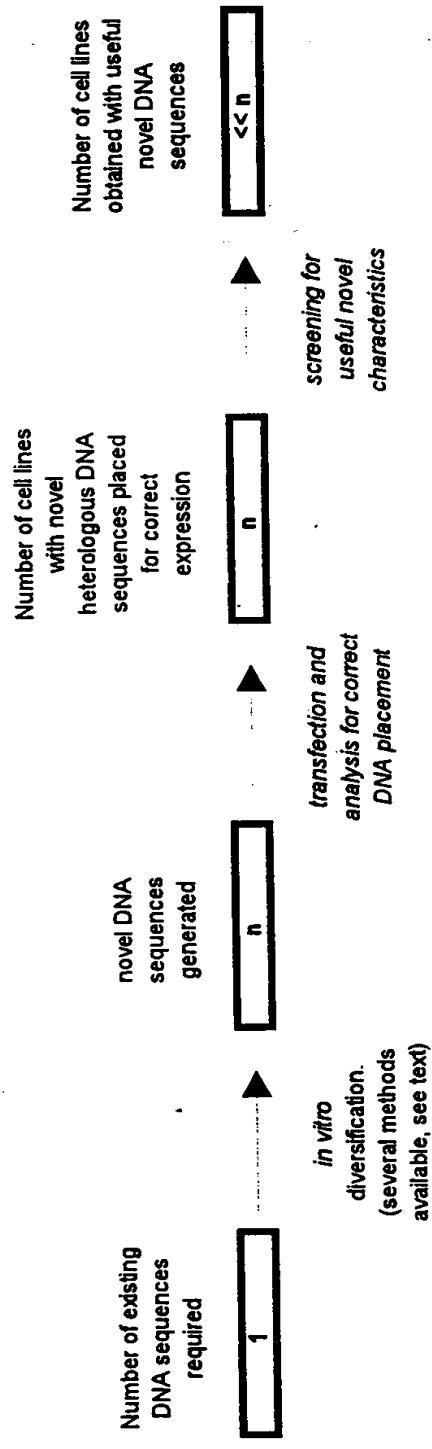


Figure 1 Methods for the diversification of DNA sequences and testing for superior variants

1A existing protocols: Number of transfections needed to generate 1024 new variants: 1024



1B a protocol enabled by the present invention: Number of transfections needed to generate 1024 new variants: 2

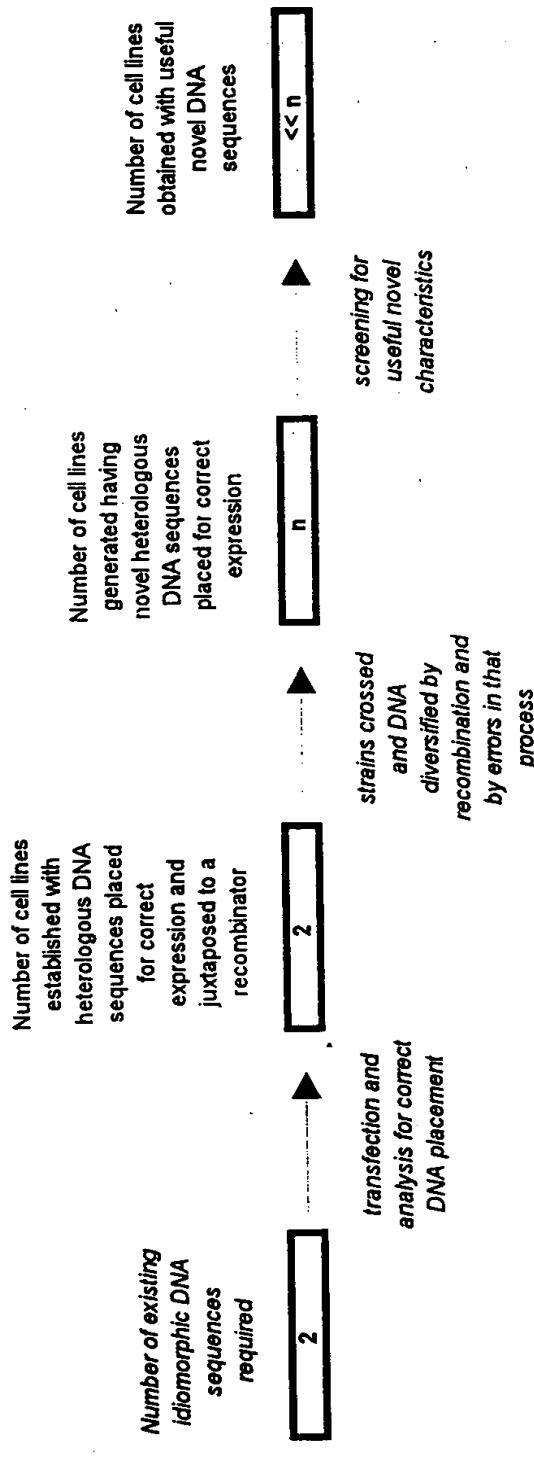


Figure 2 Methods for the diversification of DNA sequences coding subunits of heteropolymeric proteins and testing for superior variants.

The example given for immunoglobulins is for illustrative purposes only and is not intended to limit application of the present invention to this specific heteromeric protein. H = heavy chain genes, L = light chain genes

Figure 2A existing protocol: Number of transfections needed to generate 1024 new combinations: 2048

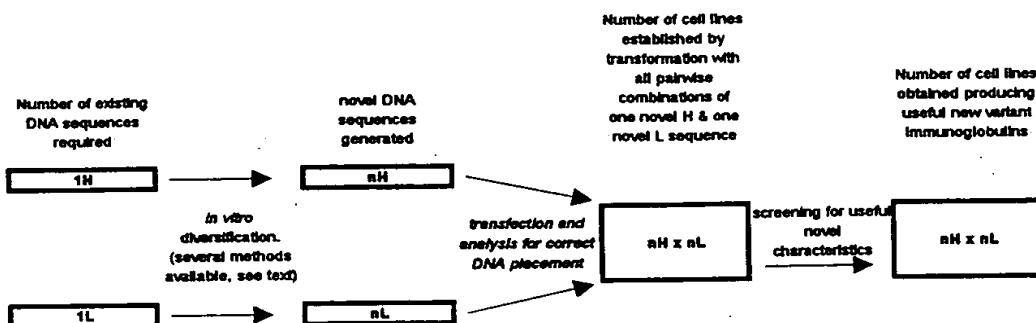


Figure 2B existing protocol using the heterokaryon technology of US Patent Serial No. 5,643,745
Number of transfections needed to generate 1024 new combinations: 64

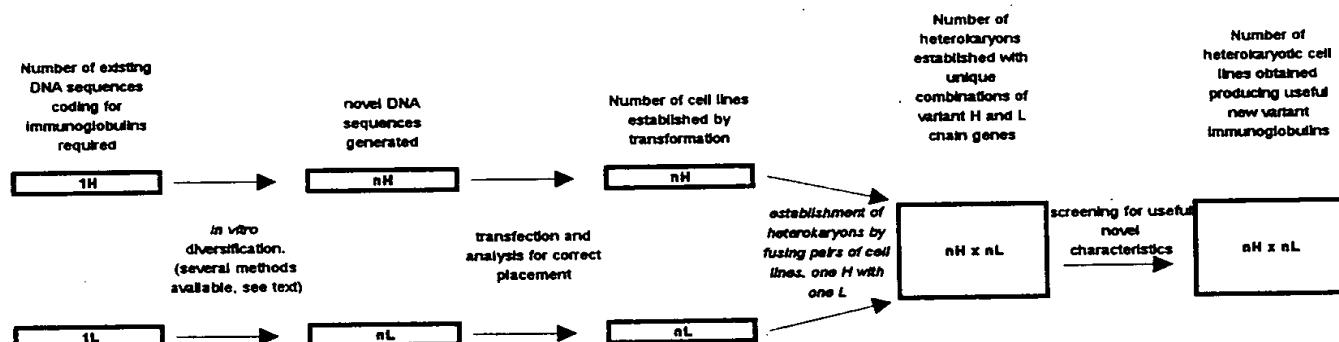


Figure 2C a protocol enabled by the present invention
Number of transfections needed to generate 1024 new combinations: 4

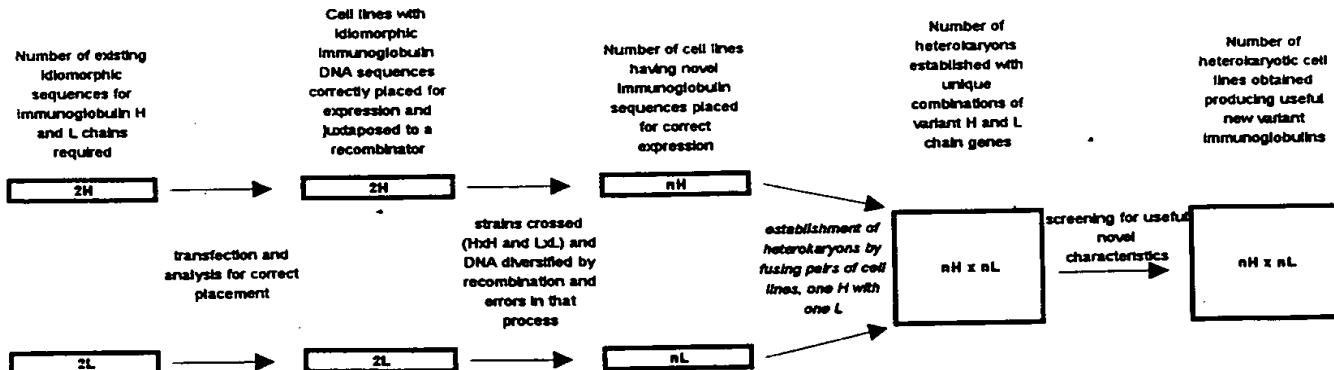


Figure 3 The modified double strand break repair model for meiotic recombination. After H Sun *et al* Cell 64: 1155-1161, 1991

(a) A double strand break (DSB) is made in one DNA duplex. (b) A long 3' overhanging single strand tail is generated either side of the break by resection. (c) One 3' end invades a homologous duplex forming a D loop. (d) the D loop is enlarged by repair synthesis and anneals to the second 3' end. (e) Repair synthesis occurs at the second 3' end and two intermolecular junctions (Holliday junctions) are formed. Resolution of the junctions by cutting inner and outer strands can give rise to non-crossover (f) and crossover (g) chromosomes. If there are base mismatches in the heteroduplex regions (duplex molecules with thick and thin lines) there will be gene conversion. If mismatch repair does not occur there will be post meiotic segregation of new sequence combinations.

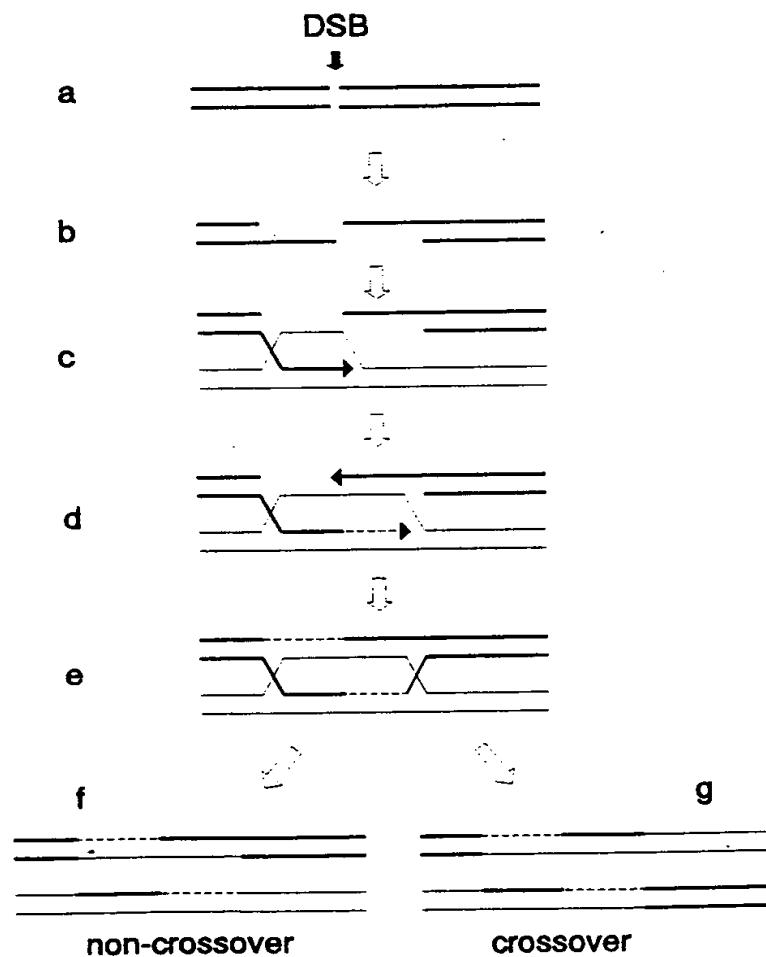


Figure 4 Life cycle of *Neurospora crassa* after JRS Fincham (Genetics, Wright 1983). Microconidia having one nucleus are not shown but can be generated as described in the text. Perithecia and protoperithecia are shown in section.

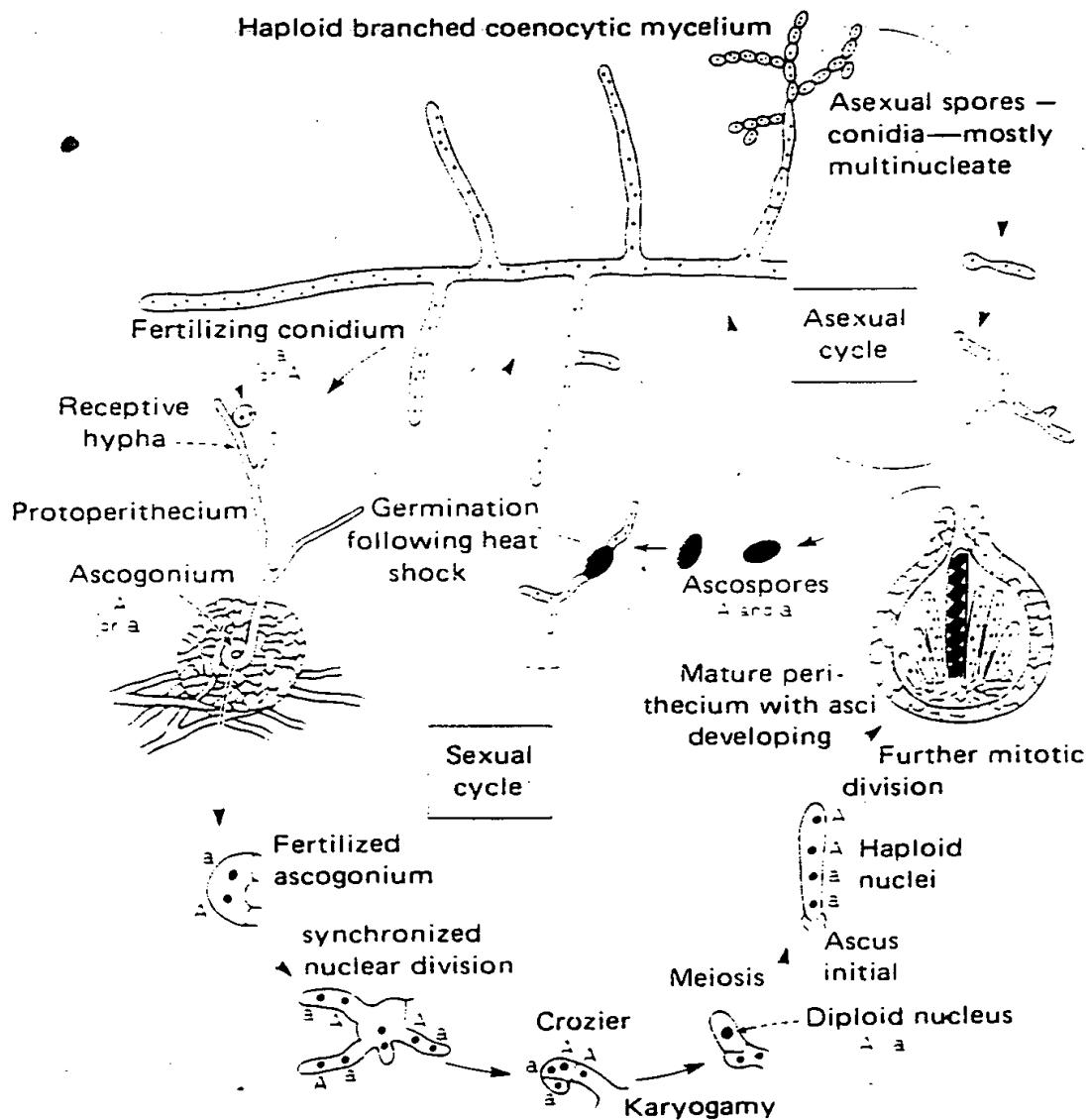


Figure 5 Map of the *his-3*, *cog*, *lpl* region of Linkage Group I of *Neurospora crassa*. Vertical bars, triangles and hairpins show the location of sequence differences that distinguish the St Lawrence and Lindegren wild type strains. The corresponding full DNA sequences are given in figure 7 and figure 8. Vertical slashes indicate one to seven base substitutions per 10 base pairs. Triangles indicate short sequence insertions and the hairpin a 101 base pair inverted repeat transposon fragment present in St Lawrence.

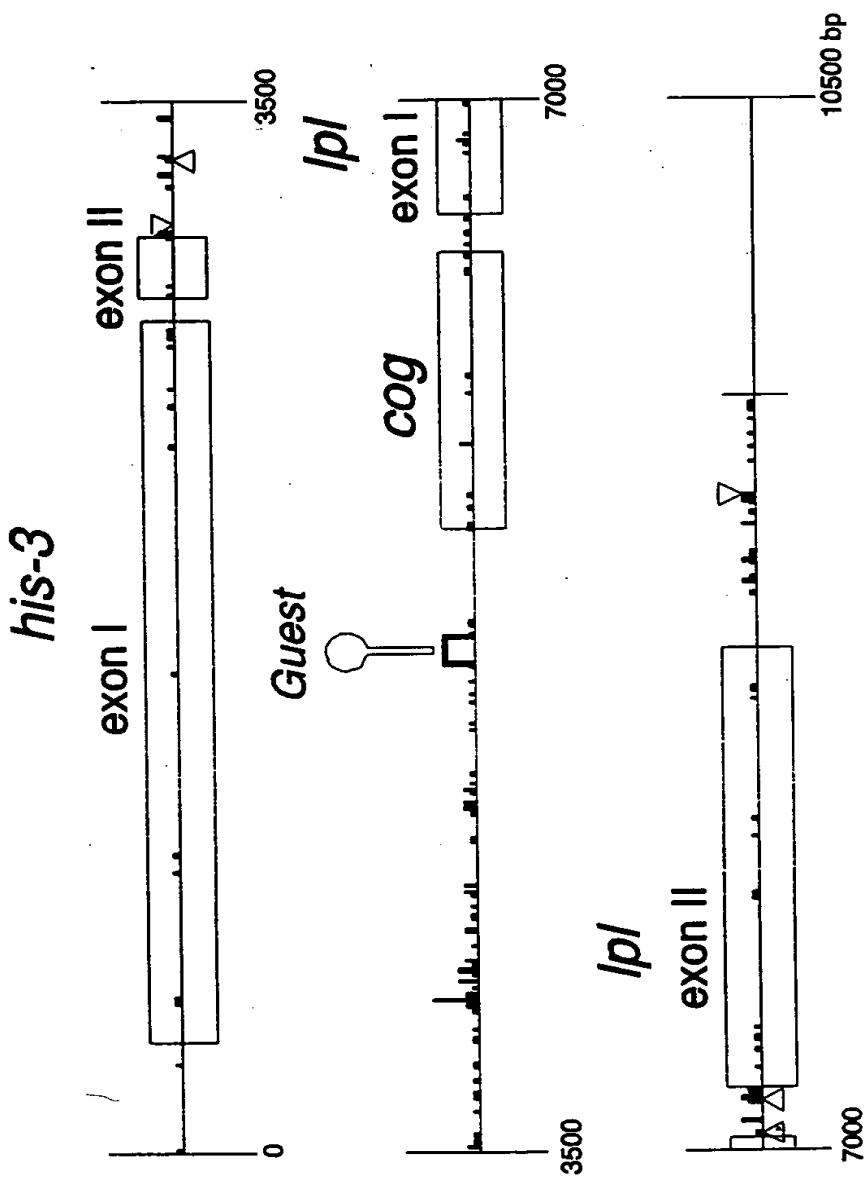


Figure 6. Discontinuity in the parental origin of DNA sequences in progeny from crosses between pairs of *his-3* alleles. In most cases this reflects discontinuity of conversion tracts, in some cases crossovers near the ends of conversion tracts.

Markers are specific DNA sequence differences that distinguish the parents. These were all E (Emerson wild type origin) or all L (Lindegren wild type origin) in the parental strains. Recombinants carry both E and L markers.

Marker position is given in base pairs from the first base of the first codon of the *his-3* gene. Each line of the table shows the parental origin of the markers inherited by one of the progeny.

marker	P	H	P1	K1201	K504	L3	R1	K26	K874	R4	C4	C5	C6	C6/7	C7	C8	C1	C2	C3	C9	D
location	~600	-384	115	179	563	1232	1502	1717	2318	3436	3705	4000	4304	4667	4821	5232	5495	6153	6507		

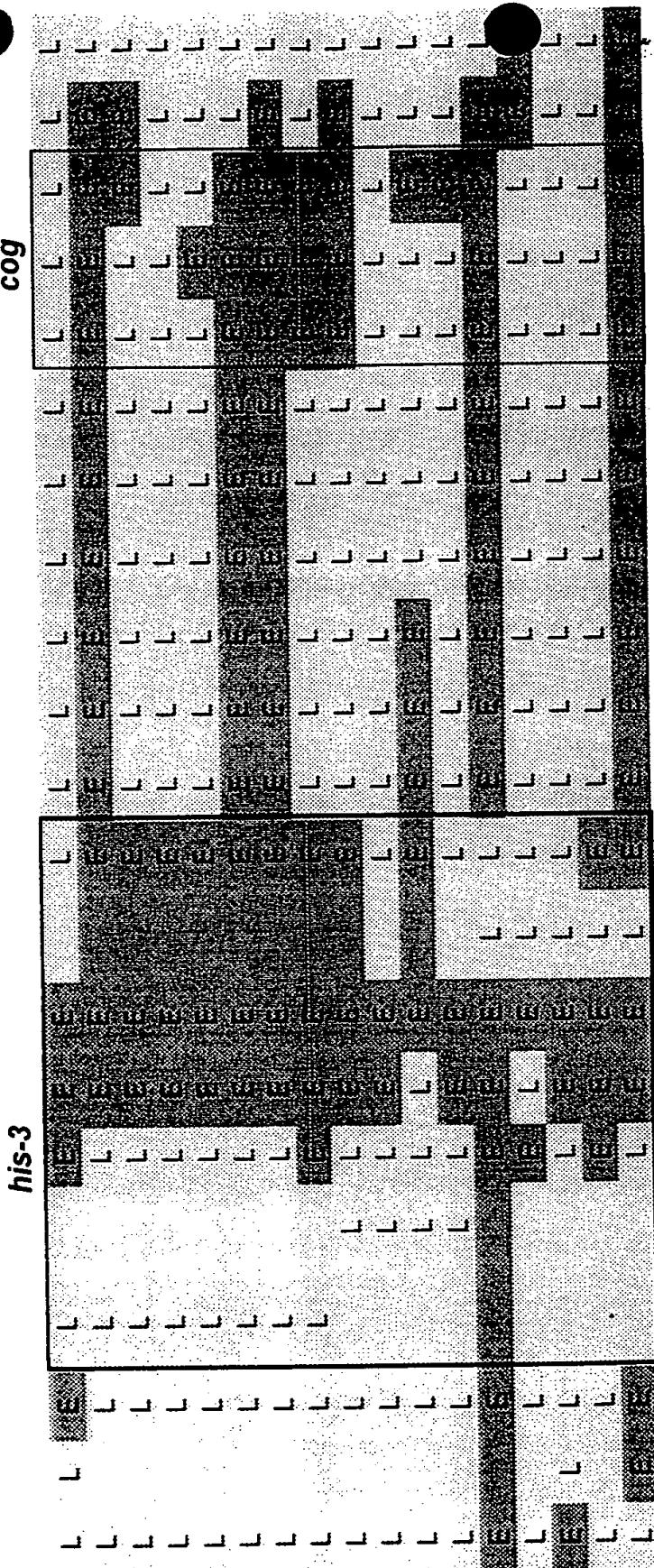


Figure 7 Nucleotide sequence of the *his-3 cog^L lpl* region of linkage group I in the Lindegren wild type strain of *Neurospora crassa*. This differs from that in the StLawrence strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the high frequency recombinator *cog^L* which is active providing the cross in which meiosis occurs is homozygous *rec-2*.

1 GATCGCAACT GGAGATCACT CGCACCGTGC CGCAGAACAA GGGCGACGAG CCTCAGGGCA
 61 GTTTAGCCTG CCGTAACAGC ACAGACCATA GCTTATTTC ACCTGGGCGG CGGGCGACG
 121 GCGGCACCTGA CATCGGCAAG GCGGCATCAA GCAACCCCTC TGTTGCTTGC CAGCTGCCGG
 181 CCAACGTCAG CGGTACAAGG AGAAATCTGG AAGGAAAGAC TTCTGGCACC GACAGGATGG
 241 CACGCGGGAA AAGTTCCCAA TGCATGAGAT GAGGGGCATT TGCATTGCCT CCCGTCACAC
 301 TGCCCCGAA CCCAACCCCC ACCATAGCGT CTGTCGATAC ATGGAGCGCG AAGTCGAGAA
 361 ACCTGTAATT CCTGGTAACT TTCAGGTACA CAGTACGTAC TGATCCTGGT ATCAAAACCTT
 421 GCCTGCCGAG TTTTCGACGG AAAGAGGTGT GAATTGTGAA AGAGTCATAC CAAATCACCC
 481 GATTTCATA AAGCCCAGT CTTTCTGTA CATAAGCGAC ACTCGAAGCG GGCCTCATCT
 541 TCATAGCCTG ATAGCTTGT AATACCCATC CTCGTATCTC ACTTGACCTT GAGTCAACC
 601 CCACGTCAGA CTTCACCCGA CACATCGACG GATTGGGAA CAGCACAATA CCTGAAAAGC
 661 GAGAAAACCA AACAGAGGAA AACACCATGG AGACAACACT TCCCCTCCCC TTCCTCGTCG
 721 GTGTCACTGT TCCCTCCCGA CTGAATGACA TCAAGGAGGG CCTCAGCCGG GAGGAAGTCT
 781 CGTGTCTTGG CTGCGTCTTC TTCGAGGTCA AGCCCAAGAC CTTGAGAAA ATCGTGCAT
 841 TCCTCAAGCG TCACAATGTC GAATTGAGC CCTACTTCGA TGTAACAGCC CTCGAGTCTA
 901 TCGATGATAT TATCACTCTT CTGGACGCC GCGCCCGAA GGTGTTTGT AAGACCGAGC
 961 AGTTGGCCGA CCTCTCCGCA TATGGCTCCC GCGTTGCCCG CATTGTCACT GGAAGCAGCG
 1021 CTGCTTTGCT TTCCCTCCGCC ACCGAGAGCC GCCTTTGCT CTCCGGCTTC GATCAGACTG
 1081 CCTCCGAGGC TGACAGCTT CTGGAGGGAGG CCAGAGACAA GAAAATTACCC CCGCTTCTCA
 1141 TCAAGCCCGT TCCCTGGGCC GATCTCGAAC AGTTCATCCA GGTGCGCCGG AAGGCTAACG
 1201 CCATCCCCAT CCTGCCATCC ACTGGCTTGA CAACAAAGAA GGACGAGGGC GGAAAGCTTG
 1261 CCATCTCCAC CATCCTCTCG AGCGTCTGGA AGTCTGACCG TCCCGATGGT CTGCTCCCCA
 1321 CCGTTGTCGT TGATGAGCAC GACACTGCTC TGGGCTCTGGT CTACAGCAGT GCGGAGAGTG
 1381 TGAACGAGGC CCTCAGGACA CAGACTGGTG TCTATCAGAG CCGGAAGCGC GGTCTCTGGT
 1441 ACAAGGGTGC TACTTCCCGA GACACTCAGG AGCTCGTCCG CATCTCGCTT GACTGCGATA
 1501 ACGATGCTCT CAAGTTTGTCA GTGAAGCAGA AGGGTCGTTT CTGCCACCTC GATCAGTCCG
 1561 GCTGCTTTGG TCAGCTCAA GCCCTTCCCA AGCTCGAGCA GACTTTGATT TCGAGGAAAC
 1621 AGTCTGCCCG CGAGGGCTCC TACACTGCC GTCTCTTCTC CGATGAGAAAG CTAGTCCGGG
 1681 CCAAGATCAT GGAGGAGGCT GAGGAGCTCT GCACCGCTCA GACCCCCCAG GAAATCGCT
 1741 TTGAGGCTGC CGATCTCTTC TACTTTGCTC TTACCAAGGGC CGTTGCTGCC GGCCTTACTC
 1801 TTGCCATAT CGAAAGGAGC CTTGACGCCA AGAGCTGGAA GGTCAAGCGC AGGACTGGAG
 1861 ATGCTAAGGG TAAGTGGCT GAGAAGGAGG GCATCAAGCC TGCGCGTCC GCTCCCGCTG
 1921 CCACCTCGGC CCCTGTCAACC AAGGAGGCC CCCAGGAGAC CACCCCTGAG AAGATCACCA
 1981 TGAGACGTTT CGACGCCCTCC AAGGTCTCTA CCGAGGAGCT CGATGCTGCT CTCAAGCGTC
 2041 CTGCGCAAAA GTCGTCCGAT GCCATCTACA AGATCATTGT CCCCACATCATC GAGGACGTCC
 2101 GCAAGAACGG CGACAAGGCT GTTCTGTCGT ACACCTCACAA GTTCGAGAAC GCTACCTCTC
 2161 TTACTAGCCC CGTCCTGAAG GCGCCCTTC CCAAGGAGCT TATGCAGCTC CCTGAGGAGA
 2221 CCATTGCTGC CATCGACGTG TCCCTCGAGA ACATCCGCAA GTTCCACGCC GCCAGAAGG
 2281 AGGAGAACCC CCTCCAGGTC GAGACCATGTC CGGGTGTGTT CTGAGCCGT TTCTCTCGTC
 2341 CCATCGAGGC CGTCGGCTGC TACATCCCCG GCGGTACCGC CGTTCTCCCC AGCACTGCC
 2401 TTATGCTGGG TGTTCGGCC ATGGTGCCTG GCTGCAACAA GATTGTGTTG GCCTCTCCTC
 2461 CCCCGGCCGA CGGAACCATC ACTCCCAGAGA TTGTCCACGT CGCTCACAAAG GTGGGGCCG
 2521 AGTCATCGT GCTTGGCCGG GGTGCCCAGG CGTAGCTGC CATGGCCTAC GGCACCGAGA
 2581 GCATCACCAA GGTCGACAAG ATTCTCGGCC CGGGTAACCA GTTCGTCACT GCTGCCAAGA
 2641 TGTCGTCAG CAACGACACC AACGCTGCCG TTGGGATTGA CATGCCCGT GGCCCGTCCG
 2701 AGGTGCTGGT CATCGCTGAC AAGGACGCCA ACCCCCGCTT CGTTGCTCTG GATCTCTGT
 2761 CCCAGGCTGA GCACGGCGTT GACAGTCAGG TCATCTCTGAT CGCTATTAAC CTCGACGAGG
 2821 AGCATCTCA GGCTATTGAG GACGAGGTTC ACCGTCAGGC TATGGAGCTT CCTCGCGTCC
 2881 AGATTGTCGG TGGCTCCATC GCCCCACTCGA TCACCGTGCA GGTCAAGACC GTCGAGGAGG
 2941 CCATGGAGCT CAGCAACAAAG TACGCTCCTG AGCACCTGAT CCTCCAGATC AAGGAGGCCG
 3001 AGAAAGCTGT CGATCTGTC ATGAACGCTG GTAGTGTCTT CATTGGCGCT TGGACTCCTG
 3061 AGTCCGTTGG CGATTACTCT GCTGGTGTAA ACCACTCGCT GCGTAAGTTA CATATCATAA
 3121 ATAGCCCCGC TTCACAGATT CTTCTGCTAA CGTCAAGACA CATAGCTACC TATGGTTTTG
 3181 GCAAGCAGTA CTCTGGCGTC AATCTCGCCT CGTTCGTCAA GCACATTAC AGCTCCAACCT
 3241 TGACTGCCGA GGGTCTAAA AACGTCGGCC AGGCTGTAT GCAGTTGGCT AAGGTTGAGG
 3301 AGCTCGAGGC TCACAGAAGG GCGGTACGCC TCCGTCTTGA GCACATGAGC AAGAGCAACT

Figure 7 continued

3361 AGACGGAAAT TCTTTTCGA AGTTGCAAAA AAAACAAGAA CAAAAGGATG TAGTGGTTG
 3421 ATGTATATCT GGGTCATTG GGGCACATAG AGTAATGATA ACGAGTTTG GACATTGTAC
 3481 TGTTCTGTAC AGGCTGAAGA TCAGTACATG AATCTGTTGG TAAGTGTAGA GACCCAAACG
 3541 TCCCCTTGAGT TTTTCTCCCT GTTCCAGAGA GGTGCTCGTC CCTGGGTGTT TATTTTCATT
 3601 ATTACATCAA CCTTTTATTT TATTTTATTT TTTATTTTAC TTTTTTTTCC TTTTTTTTCAG
 3661 ATCATCGTA CATGAACGGG GGAAGCACAG ACGATCGAAA CGTGGATGTC ACAATGTCGC
 3721 TGCAGTGATG CTGCATTGCA TGAAGCGCCC ATCTCAATAT ACTTGCAGTC TTGCGCGTTG
 3781 CACGTGAACT TCCCCAACAA CCGAATAAAA GACGGCGAAA ATGAAGATA AAAAACC
 3841 ATAATAAAAA TCGGAGGGAG TGTTGGAAAT GTTTCTTTT ACATTTAGA CCCCATAGCC
 3901 GTGCACGCCG GGGTACAGAC AGGTTCATCG ATGTTGACAT TGACTGGGAC ACCAGGTCTA
 3961 TCTATTCAT CTCCCTGCTCT CTACCATACA TCGGGACATC GGACATCTCG CTGACCCCC
 4021 CACACCCACA AAGTCTTATA AAAGCGCCAC ACCCGAGGAG GTTCGGTCCG CCCCACGAAAC
 4081 TCCGTGCTC CCTGCTGTT TACAGGGACC GAACGCTGGA GAAGCTTAGT TTCCTGACAT
 4141 CGGGCCTACC CGAGCAGGAA AAGGGACAGC TCATAGGCGA GGAGGGATTG GAAGATGGGG
 4201 ACATTTGGA TGATTCGAGA GGAGGAACTA GGTACTGTAT CATGATAGTT CGGGCAGCA
 4261 TCTTGGCTGG GACATTGTTA ATACCTCGAT ATGATGAAGT GGGAGGGAGT TTTTTCATGT
 4321 CTTGCCAAG TCCCCTAAT CTTTTTTTTT TTTTGTACCA ACACCCAAGA TTGGAGGAAT
 4381 AGTGTAAAGGA TTCGCATTCA CAAGTGGAAAG TCTGAGGATC TTTTATATC TTGCTTCC
 4441 GCGGACTGTT AACGATCTA CAGCGAGCGA GCGAGCGGTG GGATGCGCTG ATCTGATAGG
 4501 TGCAATATAC GGCCGCTTC TCCCGTCTG TAGTGTAAAG TCTGCGGCA TAGTGTAC
 4561 CTAAAAAAAC CTTGCATTG CATGATCTGC TTGCTATTCA TTGGAGTTA TTCACTGTTG
 4621 CACATTCGA GATTACAGC CATCCATCCA TATGGAAAAA TCCATTCCA TGCTTCC
 4681 CCCCCACTAT GTATGTGACC ACACGCTGCT GTCAGAAATGC CAACGGCTC AGGTACCC
 4741 GTCCGACTGT TTGGCATGGA GTTACATACA CTACTAGTGT AGCCCCGGGC CAAGCTACCC
 4801 CGTCAAATCT ATACATATCT ATAATGGGTT TCAGGTGTT CGTTCGCTGT CAATCAAGTT
 4861 TGAAACATCA CTGGGGCCGT TGGACGGTGT ATTAGACCAT TGGCTCCCTC AGCTGGCGC
 4921 TGGCGGTTG GGTCGGCAAT AACGGGACTG GACTTGAGAG GGACGAGGAG AGTCGGTTG
 4981 CTGCCCTACAC TACACTACAA GCGTTCACAC CTAACCGACG AGTCCCGTT TCCATTGTC
 5041 TGCCTTAACC ATCATCTAGG GATGTCAAGG TTTGGCCGA TCAGGGTATG TTTGGTTGAC
 5101 TGTTGTCTG TCTGATTGGG TACATATCAT GGTAGGTGTC TCGAGAACAG TAGACTCTC
 5161 GGGCTAGCG TTTGGATGAT TACCGCAGAT ATGAGTTGTA GGCCGCCATG CAGTTGCTTG
 5221 CCCATAAGCA GGAGTTGCTT TGGGATATAT TTCTCGTCTT TCAAAGGTCA CGAGGTCTG
 5281 GGACGAGCGG CATGCCATC CAAAGGGTTG AACATGAGAA ACCGGAATGG CCTTGCCTT
 5341 GAAATACAAA AAGTCAAGAA TAAAATCGCT TGAGGATAGG GACGTGGAAG CAAGCAAATA
 5401 TGGTAAGGGA GGTACTGCTA TGAGGTGCT CAGCAAACCTG CCAATTCTT GGCCCCAAG
 5461 CAGCAGTTG CTGTCAGTGC TGCTCGTGT AGCCTGGTA GTGGAACCTA AACTGCTAAC
 5521 ACAGCGCAAG TGCGCATGTA AAGATATTGT GGGAGGATCT TATGGATGG ATGAGGATAC
 5581 TGCTTGGTGT TGGTTGCAG GCACTGCGG TGTTAGGCTT TGCTGTGCC CGTTCGACGA
 5641 AGAAATACGC GGAACATAA ATTGGATACC TAGACTTACT GCCTATGGGA GGTATCTACC
 5701 GACGTAGCGG ACGGATTCTA GCAACATCCC GACTTGTCTT GTAGTGTACT ATGATAGCAG
 5761 CACAGTGGGG TGTGCTCT TGTGAGCATG GGCTCTTTT TTTTTTTTCC CCCTTCCCTA
 5821 GGGCGTTGAC TGGACTTGCT CTATGTTCC CAAGGTAGGT GCCCGTCATC GATTTTCCCA
 5881 AGCCGCTCC CGCCAGATTG TCGTCATAGT GTCATGATGA CCTCGGTCGC TGGGCTGCG
 5941 TGGTTACGGG GAGCTGGAC CGCTAGGCC CAGTGGTTGT GCCATTCAAG GTGGGTGTGT
 6001 GGAGTAGCGG TAGAGGCCT TGAAGTTGT GCTAGCGAA ACCCTGGAAT ATCTTGTACC
 6061 CTTCGATTCC TTCTCGGGCT GCCCATGTGC TGAGGTGATG CCGGGGATCT GCGCCAATC
 6121 ATCCATTGAG GTTCCCGCAG CTCCCGGGT CCGCGCCCGG GCGCAGTTGC TCACAGGACA
 6181 CACCTAGACG CAGGGGCACA GGGGCACCGT TTGGTGTCA ACTGGGTACC TGGTAGTGT
 6241 AGCAAGCACT CCACCGCTG TGCATATCCC CAATCCACCG CAGGAACCTA GCACCGCCGC
 6301 GGCACCGAGT GAGCGAATCC ATCCGCATTG GATCCCAATT CTTGCCCTG CCATCCTTCT
 6361 TTCTTCCAC TTGGCGAAC CAACACTTCC TTGGTCTGG GTACTCGTGT TGATCTTAC
 6421 TCTCTTTTT TCTTGGCGA CCGACTTTT ATATCCGTCC TTGCTTCCCC CTGGCCGTTG
 6481 TCGTTCTTC TACAACCTAC TTCCGTTCAT TATCCCTTT CTTGGTTCGG TCGAGGACCC
 6541 AAAAACAGAA CAATTCCGGC TCTTCCAGGT GGCTTGGGT CGACTGTTA GCTCTGACC
 6601 ACTAGCGCT TACCTTCTCT TGATGTTAT ATTTGGATAT CATTGAACTA CTCTTCTTG
 6661 AAACGGCAGA CGAACGGAAC AGTCCCTACG GTTATTAGC GATATACGTT GTACTGATAT
 6721 CCTGAGCAAG AAGAGGCAAAT TTATCAATT TGCACTCTCC ATCGTCGCTG CTCATCGCAG
 6781 CTCCCTTGCT CGCCAATGTA TCGGCGAAC CGATTAGGAT ACCCCAAACGC GATGTTCTCC
 6841 GTGGTATCAA CATCACAGCA ACTGCCGT CGAGCACTAC CGAATTCGCC CAGCGGTGGA
 6901 TATGCCCTG CGGTTGAGA CTGTCCTAAC ACCAAGCCGA CGCTCCGGAA GGCGTGGAT
 6961 TTGTCGAACG AGGAGAAGAA CTGGTTGTCG ATCCGGAGGA AGAACACCAT CCAGCCCATG
 7021 AGGGACCTAC TGAAGAGGGC CAACATCACT GGGTTGATT CGAAACTTT CATGAATGAG
 7081 GCCGCCAACA ACGTCTCGCA ACTGCCCAAT GTCGCCATT CCATTTCAAG AGGCGGCTAT
 7141 CGTCCCTCA TGAACGGCGC CGGCTTCGTT GCTGCTGCGG ATAACCGGAT TCAAAATACC
 7201 ACGGGCGCAG GTGGTATTGG AGGCTTGTG CAGTCCAGCA CATATTGTA TGAAAACCA
 7261 TGCCTTCTTG TGGTTCTTCT TATCTCGTT TCGAGTGTCA ACTGCGCCAG TTGACGCTTG

Figure 7 continued

7321 GCGGGCTGTG GACGACCTTG CTGGTGAACA TGTCTGGAC TCCATGCC CTTTTTCGTT
7381 CCCTAAAATC CCAAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAATTGAG
7441 GACCGTGA CTAATGGCT AACGCAACTC TAGGGCCGGA CTTTCTGGTG GTGGCTGGCT
7501 TGTGGCAGT TTGTTCTCCA ACAACTTCAG TAGCATTGAG ACCCTGCTGA GCGAGAACAA
7561 AGTCTGGAC TTTGAGAACT CCATCTTAA AGGACCCAAG GAGGCTGGCC TTAGTACTGT
7621 CAACCGTATC CAGTACTGGT CGGAAGTGGC AAAGGAAGTT GCGAAGAAGA AGGATGCTGG
7681 CTTCGAGACA AGTATAACAG ACTACTGGGG CCGAGCATTG AGTTACCAAC TGATCGGAGC
7741 CGATATGGC GGCCCGGCTT ACACCTTCTC CAGCATTGCC CAGACCGACA ACTTCCAGAA
7801 GGCGGAAACG CCGTTCCCTA TTCTGGTAGC TGACGCCGC GGCCTGGAG ACACCATCAT
7861 CTCCCTCAAT GCTACCAACT ACGAGTTCAA CCCGTTCTGAG ACGGGTAGCT GGGACCCGAC
7921 CGTCTATGGC TTGCGCCGA CCAAGTACCT CGGCCAAC TTCAGCAACG GCGTGTACCC
7981 ATCAGGAGGC AAGTGCCTTG AGGGTCTCGA CCAAGCCGGC TTCTGTATGG GCACCAGCAG
8041 CACGCTCTTC AACCAGTTC CTTTGGCCAA CATCTCCAGC TACGACGGTG TTGCCAGACG
8101 TGCTCATCGA GGCGGTGACT TCTGTCCTCA AGGAAATCGG CGCCAAGAGG ACGACGTCTC
8161 CCAAATCATC CCTAATCCGT CCTGGACTG GAACAACCG ACCAACCCCA ACGCCGACAC
8221 GCTCGAGCTC GACCTGGTCG ACGGCGGCCA AGATCTGCAG ATATATCCGC TCAACCCGCT
8281 CACCCAACCC GTGCGGCCG TCGACGTCAT TTCTCGTGT GACTCGTCCG CCGACGTGAC
8341 AAACTGGCCC AATGGCACCG CCCTGCGCGC CACCTACGAG CGCACTTTCG GCTCTATTTC
8401 CAACGGGACA CTCTTCCCCCT CGATCCCCGA CGACTGGACG TTATTAACCC TAGGCCCTCAA
8461 CAACCGCCCC TCTTTCTTCG GCTCGATGT TAAGAACATT ACCTTGAACG CCAACAAAAA
8521 GGTTCCCCCC TTAATCGTCT ATGTCCCCAA CGCGCCCTAT ACCGCGCTGA GCAACGTGTC
8581 CACCTTCGAT CCGTCATACA CGATGTCTCA GCGCAACGAC ATCATCGGCA ACGGATGGAA
8641 CTCAGCCACG CAGGGAAACG GCAAGCTGGA TTGGAGTGG CCCACTTGC CGCCTGCC
8701 GGTTATCAGC AGGAGCTTAG ATCGGTTGGG CAGGCAGACG CCAGCCGCGT GCAAGACTTG
8761 CTTTGACAGG TATTGCTGGA ATGGCACAGT GAAACTCCAAA GATACGGGGG TTTACATGCC
8821 TGAGTTCAAG ATTGCGGATG CGCATGCCCT GGACTCGGGT GCTGTTGCTA TCGGAAAGAT
8881 GGTGAATGTC TGGTCGTGGG TTGTGGTGGG AGTTGTGGCG GCTACTTTGT TGTGTAGGG
8941 GTAGGGGAGA CGTGTGATA TTCCAGTCTG ATGAAGTTGA GACTGGACTG GAGATCGCCA
9001 AGGATGCGGA GGGAAAGGAA TCGGTGGTGT TAATGTCATG ATGGATGAAG AGTCATGGAT
9061 CATGAAACGA CGGGCGGGG ATATTGGATG ATGGATATAC CACACTGCAT GCATGCTCTA
9121 TTGATAGTAT GCTTTGGCAT TTACTTTAA CAATCAATTG CTCCATCTG ATGTTCTATC
9181 TTTTCGACA ATGGATTGAT ACTACTCTG TTGCTCGCT CTTGAGGTG GAAGGACTTG
9241 AGGTTGGAAG GACTTGAGGT TGTGGTTCT GAGGGAGGTT ATCGAAGTAT CATCTGTGCT
9301 GATGCCGATT GATAGACTGT CCTCTTCTTC GAGGCAACGA ACGGTCGGAT GAGCCTCTT
9361 AATCATGATG CTCAGTGCCA CAAAAAGGCT CCAGCACAGC TGCCCACACC TTTCTGCCT
9421 CGCCGTTCCCT TCCTTTTCT TTTCCCTGT TTCTTTCTT CCTTCCATC TCATCCCCTA
9481 CCAGAGTGC CACCGGGTAT ATATATTAC TCCCTGGCCG TTCTCTTTG ACCAATAAAAT
9541 CGCTTGGTCG AGTGGCGTAA CGGTTTACCG TCTACACTTA TCACTCAAAC CAAACCAAAC
9601 CATCGAAGAA GTGACCTATC GGTTCGAGGG AACGGTGATG TTCTTACGAC CAAGTTAAC
9661 CAAAGAGCGT TCCACATCGT TGAACCGTCT CCTCCAGTTG GATCTGTTA ACTCCCGCAG
9721 CGACTGAAGA AGGTATCACT TTTTTTTGG TTCCAAAAAA AAAAAAAA ATTAC

Figure 8 Nucleotide sequence of the *his-3 cog^E lpl* region of linkage group I in the StLawrence wild type strain of *Neurospora crassa*. This differs from that in the Lindegren strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the weak recombinator *cog^E* and also the remnant of a transposable element *Guest* within the replaceable sequence 3' of *his-3*. StLawrence strains carry *rec-2⁺* which prevents the initiation of recombination at *cog*.

```

1 ACCGGGAATC GTAGCGGGCG CTAAGGCCAA GCGCGGCAC GGGTCACTGA CCCAATGCG
61 CGCATTCCGT CAGCACTGA AGTGGATGTA CAAGTACATA GTAGTAGATC GCAACTGGAG
121 ATCACTCGCA CCGTGCAGCA GAACAAGGGC GACGAGCCTC AGGGCAGTT AGCCCTGCCGT
181 AACAGCACAG ACCATAGCTT ATTTCACCT GGGCGGGCG GCGACGGCGG CACTGACATC
241 GGCAAGGGG CATCAAGCAA CCCCTCTGTT GCTTGGCAGC TGCCGGCCA CGTCAGCGGT
301 ACAAGGAGAA ATCTGGAAGG AAAAGACTCT GGCACCGACA GGATGGCACG CGGGAAAAGT
361 TCCAATGCA TGAGATGAGG GGCATTTGCA TTGCCTCCCG TCACCCAGTG CGAACCCCCAA
421 CCCCACATCA GCGCTGTGCG ATACATGGAG CGCGAAGTCG AGAAAACCTGT AATTCTGGT
481 AACTTCAGG TACACAGTAC GTACTGATCC TGGTATCAAA CCTTGCCCTGC CGAGTTTCG
541 ACGGAAAGAG GTGTGAATTG TGAAAGAGTC ATACCAAATC ACCCGATTTC CATAAAGCCC
601 GAGTCTTTTC TGTACATAAG CGACACTCGA AGCGGGCCTC ATCTTCATAG CCTGATAGCT
661 TGTAATACTC CATCCTCGTA TCTCACTTGA CCTTGAGTTC AACCCCACGT CAAACTTCAC
721 CCGACACATC GACGGATTGG GGAACAGCAC AATACTGAA AAGCAGAGAAA ACCAAACAGA
781 GGAAACACCC ATGGAGACAA CACTTCCCT CCCCCTCCG GTCGGTGTCA GTGTTCTCC
841 CGGACTGAAT GACATCAAGG AGGGCCTCAG CGGGGAGGAA GTCTCGTGT TTGGCTGCGT
901 STTCTCGAG GTCAAGCCCC AGACCTTGA GAAAATCTG CGATTCTCA AGCGTCACAA
961 TGTGAATTG GAGCCCTACT TCGATGTAAAC AGCCCTCGAG TCTATCGATG ATATTATCAC
1021 TCTCTGGAC GCGGGCGCCC GCAAGGTGTT TGTCAAGACC GAGCAGTTGG CCGACCTCTC
1081 CGCATATGGC TCCCGCGTTG CCCCCATTGT CACTGGAAGC AGCGCTGCTT TGCTTCTCC
1141 CGCCACCGAG AGCGGCCCTT TGCTCTCCGG CTTCGATCAG ACTGCTCCG AGGCTGCACA
1201 GTTCTGGAG GAGGCCAGAG ACAAGAAAAT TACCCCTTC TTCATCAAGC CGGTTCTGG
1261 GGCGGATCTC GAACAGTTCA TCCAGGTGCG CGCCAAGGCT AACGCCATCC CCATCTGCC
1321 ATCCACTGGC TTGACAACAA AGAAGGACGA GGGCGGAAG CTTGCCATCT CCACCATCCT
1381 CTCGAGCGTC TGGAAAGCTG ACCGTCCCGA TGGTCTTCTC CCCACCGTTG TCAGTGTGATGA
1441 GCACGACACT GCTCTGGTC TGTTCTACAG CAGTGGCAG AGTGTGAACG AGGCCCTCAG
1501 GACACAGACT GGTGTCTATC AGAGCCGGAA GCGCGGTCTC TGGTACAAGG GTGCTACTTC
1561 CGGAGACACT CAGGAGCTCG TCCGATCTC GCTGACTGC GATAACGATG CTCTCAAGTT
1621 TGTGTGAAG CAGAAGGGTC GTTTCTGCCA CCTCGATCAG TCCGGCTGCT TTGGTCAGCT
1681 CAAAGGCCCT CCCAAGCTCG AGCAGACTTT GATTCGAGG AAACAGTCTG CCCCCGAGGG
1741 CTCCTACACT GCGCGCTCT TCTCCGATGA GAAGCTAGTC CGGGCCAAGA TCATGGAGGA
1801 GGCTGAGGAG CTCTGCACCG CTCAGACCCC CCAGGAAATC GCCTTGGAGG CTGCGATCT
1861 CTTCTACTTT GCTCTTACCA GGGCGGTTGC TGCCCGCCT ACTCTTGGCG ATATCGAAAG
1921 GAGCCTTGAC GCGCAAGAGCT GGAAGGTCAA CGCGCAGGACT GGAGATGCTA AGGTAAGTG
1981 GGCTGAGAAG GAGGGCATCA AGCCTGGGG GTCGGCTCTC GCTGCCACTT CGGGCCCTGT
2041 CACCAAGGAG GCGGCCAGG AGACCAACCCC TGAGAAGATC ACCATGAGAC GTTTCGACGC
2101 CTCCAAGGTC TCTACCGAGG AGCTCGATGC TGCTCTCAAG CGTCCTGCGC AAAAGTCGTC
2161 CGATGCCATC TACAAGATCA TTGTCCCCAT CATCGAGGAC GTCCGCAAGA ACGGCGACAA
2221 GGCTGTTCTG TCGTACACTC ACAAGTTCGA GAAGGCTACC TCTCTTACTA GCCCGCTCCT
2281 GAAGGCGCCC TTCCCCAAGG AGCTTATGCA GCTCCCTGAG GAGACATTG CTGCCATCGA
2341 CGTGCCTTC GAGAACATCC GCAAGTTCCA CGCCGCCAG AAGGAGGAGA AGCCCTCCA
2401 GGTGAGACCC ATGCCCCGTG TTGTCTGCG CCGTTCTCT CGTCCCATCG AGGCGCTCGG
2461 CTGCTACATC CCCGGCGGT CCGCCGCTT CCCCAGCACT GCCCTTATGC TGGGTGTTCC
2521 CGCCATGGTC CGCCGGTGC ACGAAGATTG GTTCCCTCT CCTCCCCCG CCGACGGAAC
2581 CATCACTCCC GAGATTGTC ACCTCGCTCA CAAGGTTGGG GCGGAGTCCA TCGTGCTTGC
2641 CGGCGGTGCA CAGGCCGTAG CTGCCATGGC CTACGGCACC GAGAGCATCA CCAAGGTCGA
2701 CAAGATTCTC GGCCCCGGTA ACCAGTTCGT CACTGCTGCC AAGATGTTCG TCAGCAACGA
2761 CACCAACGCT GCCGTTGGTA TTGACATGCC CGCTGGCCCG TCCGAGGTGC TGGTCATCGC
2821 TGACAAGGAC GCCAACCCCG CGTTCGTTGC CTCGATCTC CTGTCCTCAGG CTGAGCACGG
2881 CGTTGACAGT CAGGTCATCC TGATCGCTAT TGACCTCGAC GAGGAGCATC TTCAAGGCTAT
2941 TGAGGACGAG GTTCACCGTC AGGCTACGGA GCTTCCCTCGC GTCCAGATTG TCCGTGGCTC
3001 CATGCCACAC TCGATCACCG TGCAGGTCAA GACCGTCGAG GAGGCCATGG AGCTCAGCAA
3061 CAAAGTACGCT CCTGAGCACT TGATCCTCCA GATCAAGGAG GCGGAGAAGG CTGTCATCT
3121 TGTCTGAAAC GCGGGTAGTG TCTTCATTGG CGCCTGGACT CCGTGTGCGT TTGGCGATTA
3181 CTCTGCTGGT GTTAACCACT CGCTGCGTAA GTTACATATC ATAAATAGCC CCGCTTCACA
3241 GATTCTTCTG CTAACGTCAA GACACATAGC TACCTATGGC TTTGGCAAGC AGTACTCTGG

```

Figure 8 continued

3301	CGTCAATTTC	GCCTCGTTCG	TCAAGCACAT	TACCAGCTCC	AACTTGACTG	CCGAGGGTCT
3361	CAAAAACGTC	GGCCAGGCTG	TCATGCAGTT	GGCTAAGGTT	GAGGAGCTCG	AGGCTCACAG
3421	AAGGGCGGTC	AGCATCCGTC	TTGAGCACAT	GAGCAAGAGC	AACTAAACGG	AAATTCTTTT
3481	CGAAGTAGCA	AAAAAAAAAA	AAAAAAACAA	GAACAAAAGG	ATGTAGTGGG	TTGATGTATA
3541	TCTGGGTCAT	TTGGGCACA	TAGAGTAATG	ATAACGAGTT	TTGGACATTG	TACTGTTCTG
3601	TACAGGCTGA	AGATCAGTAC	ATGAATCTGT	TGGTAAGTGT	GGAGACCCAA	ACGTCCCTTG
3661	AGTTTTCTC	CCTATTCCAG	AGGTGCTCGT	CCCTGGGTGT	TTATTTTCAT	TATTACATCA
3721	ACCTTTTTT	TTTTTTTTT	TTTTTCAGAT	CATGCGTACA	TGAACGGGGG	AAGCACAGAC
3781	GATCGAAACG	TGGATGTCAC	AATGTCGCTG	CAGTGTGCT	GCATTGCATG	AAGGCCCAT
3841	CTCAATATAC	TTGCAGTCTT	GCACGTTGCA	TGTGAACTTC	CCAAACAAACC	GAATAAAAGA
3901	CGCGAAAAAA	TGAAGATAAA	AAAAAAACAT	AAAAAAATC	AGAGGGAGTG	TGGGAAATGG
3961	TGTCTTTAG	CATTAGAGAC	CCATAGCCGT	GCACGCCCGG	GTACAGACAG	GTTCATCGAT
4021	GTTGACATTG	ACTGGGACAC	CAGGTCTATC	TATTTTATCT	CTGTCTCT	ACCATACATC
4081	GGGACATCGG	ACATCTTGT	GTACCCCCCA	CACCCACAAA	GCCTTATAAA	AGGCCACAC
4141	CCGAGGAGGT	TCGGTCGGCC	CCACGAACTC	TGTGCCCTCC	TGCCTGTTA	CAGGGACCGA
4201	ACGCTGGAGA	ATCTTACTAG	TTTCCTGACA	TCCGGCCTAC	CCGAGCAGGA	AAAGGGACAG
4261	CTCATAGGCG	AGGAGGGATT	TGAAGATGGG	AACATTG	GTGATTGAG	AGGAGGAAC
4321	AGGTACTGCA	TCATGATAGT	TCGGGGCAGC	ATCTTGGCTG	GGACATTGTT	AATACTCGA
4381	TATGATGAAG	TAGGAGGGAG	TTTTTGCCTG	TCTTGCCTGA	GTCCAGAGAT	CTGTTTATT
4441	TTATTTTTA	TGGATGTAGT	GTATCAACAC	CCAAGATTG	GAGAATAGTA	CTAGGATTG
4501	CATTACAAG	TGGAAGTCTT	GAGAATCGTT	GTATATCCTT	GTCTTCCTG	GAATGTTAAC
4561	AATCCATCAG	CGAGCGAGCG	AGCGGTGCGA	TGCGCTGATC	TGATAGGCGC	AATATACGGC
4621	CGCTTCTCC	GGTCGTGCTAG	TGTAAGCTCT	TGTTGGCATAG	TACACTAAA	AAACCCCTTG
4681	ATTTCATGAT	CTGCCTGCTA	TTCAATTCCGA	GCTATTTCAG	TGGTCACATT	TCGAGGAAGA
4741	AAGAAAGCAA	CTAAGATTCA	CAGCCATCCA	TCCATCCATA	TGGAAGAATA	ATCCATTCCC
4801	ATGTTCCCTC	CCCCCCACTA	TGTATGTGAC	CACACGCTGC	TGTAGAAATG	CCAACGGTCT
4861	CAGGTACCCCT	CGTCCGACTG	TTTGGCATGG	AGTTACATAC	ACTACTAGTG	TAGCCCCGGG
4921	CCAAGCTACC	CCGTCAAATC	TATACATATC	TATAACGGGT	TTCAGGGTT	TCGTTCGCTG
4981	TCAATCAAGT	TTGAAACATC	ACTGGGGCCG	TTGGACGGTG	TATTAGACCA	TTGGCTCCCT
5041	CAGCTGTTG	GGGGCTGGC	GGCTGGGTCA	AACGGCAATA	ACGGGACTCG	AGAGGGACGA
5101	GGAGAGTCGG	TTGGCTGGCT	GCAATACAAG	CGTCCACC	TAACCAACGA	GTCCCCTTT
5161	CCATTGTGT	GCCTAACAT	CATCTAGGGA	TGTAGGGTT	TGGCCGGATC	AGGGTATGTT
5221	TGGTGTACTG	TTGTCATGTC	TGATTGGGTA	CATATTATGG	TAGGTGTCTC	GAGAACAGTA
5281	GAGTACTCGG	GCCTAGCGTT	TGGATGATTA	CGCGAGATAT	GAGTTGTGGG	CCGGCATGCA
5341	GTTGCTTGT	CATAAGCAGA	AGTTGCTTGT	GGATATATT	CTCGTCTTC	AAAGGTACG
5401	AGGTCTGGG	ACGAACGGCA	TCGCCATCCA	AAGGGTTGAA	CATGAGAAAC	CTGAATGGCC
5461	TTTGCCTGTA	AATAACAAAAA	GTCAAGAACAA	AAATCGCTT	AGGATAGGGA	CGTGAAGCA
5521	AGCAAATATG	TAAGAGAGG	TATACATCAA	CCCTGGTTCA	ATTGTTAGCG	TGGTTCTTC
5581	TCCACGTCT	CGTTCATGAC	GGTTAACAGT	ACCAGGCTAA	CAATTAAACC	AGGGTTGATG
5641	TGTACTGATA	TGTAGGTGTC	CAGCAAAC	CCAATTCTT	TGGCCCCAAG	CAGCAGTTTG
5701	CTGTCAGTGC	TGTCGTGTC	AGCCTGGGTA	GTGGAACCTA	AACTGCTAAC	ACAGCGCAAG
5761	TGGCATGTA	AAGATATTGT	GGGAGGATCT	GTATGGATGG	ATGAGATTAC	TGCTTGGTGT
5821	TGGTGTGAG	GCACTGCGC	TGTTAGGCTT	TGCTGTGCC	CGTCGACGA	AGAAATACGC
5881	GGAACATAAA	ATTGGATACC	TAGACTTA	GCCTATGGGA	GGTATCTACC	GACGTAGCCG
5941	ACGGATTCTA	GCAACATCCC	GACTTGTCTT	GTAGTGTACT	ATGATAGCAG	CACAGTGTG
6001	CTCCTTGTGA	GAATGGGCTC	TTTTTTTTT	TCCCCCTTCC	CTAGGGCGTT	GACTGGACTT
6061	GCTCTATTGT	TCCCAAGGTA	GGTGCCTGTC	ATCGATTTC	CCAAGTCTCC	CGCCAGATTG
6121	TCGTCATAGT	GTCAATGATG	CCTCGGTGCG	TGGGGCTGCG	TGGTTACGGG	GAGCTGGGAC
6181	CGCTAGGCCT	CAGTGGTTGT	GCATTCA	GTGGGTGTGT	GGAGTAGCGG	TAGAGGCGCT
6241	TGGAAGTTGT	GCTAGCGGAA	ACCTGGAAAT	ATCTTCTACC	CTCGATTCT	TCTCGGGCTG
6301	CCCATGTCG	GAGGTGATGC	CGGGGATCTG	GCACCAATCA	TCCATTGAGG	TTCCCGCAGC
6361	TTCCCGGTG	CGCGCGCGGG	CGCAGTTGCT	CACAGGACAC	ACCTAGACGC	AGGGGCACAG
6421	GGGACCGGT	TGGTGTGCAA	CTGGGTACCT	AGCTGTAGCA	AGCACTCCAC	CGTCTGTGCA
6481	ATCCCCAAT	CCACGGCAGG	AACTTCGAC	CGCCGCGGCA	CCGAGTGAGC	GAATCCATCC
6541	GCATTGGATC	CCAATTCTTG	CCCTTGCCAT	CCTTCTTCT	TCCCACCTGG	CGCAACCAAC
6601	ACTTCCCTTG	GTCTGGTAC	TCGTGTTGAT	CTTCACCTCTC	TTTTTTCTT	GGGCGACCGA
6661	CTTTTATAT	CCGTCTTGC	TTCCCCCTGG	CCGTTGTCGT	TCTTTCTACA	ACTACCTTC
6721	GTTCAATTATC	CCCTTCTTG	GTTCGGTCGA	GGACCCAAA	ACAGAACAAAT	TCCGGCTT
6781	CCAGGTGGCT	TGGGTGCGAC	TGTTTAGCTC	TTGACCACTA	GCCGCTTAC	TTCTCTTGAT
6841	GTTTTTATT	GGATATCATT	AAACTACTCT	TTCTTGAAAC	GGCAGACGAA	CGGAACAGTT
6901	CCTACGGTAT	ATTAGCGATA	TACGTTGTCAC	TGATATTCTG	AGCAAGAAGA	GGCAAATTAT
6961	CAATTATGCA	TCTCCCTTCG	TCGCTGCTCA	TCGCACTCC	CTTGCTCGCC	AATGTATCGG
7021	CCGAACCCAT	TAGGATACCC	CAACGCGATG	TTCTCCGTGG	TATCAACATC	ACAGCAACTT
7081	GCCGTTGAG	CACTACCGGA	TTCGCCCA	GGTGGATATG	CCCCTGCGCT	TGTAGACTGT
7141	CCCAAGACCA	AGCCGACGCT	CCGGAAGGCC	GTGGATTGTT	CGAACGAGGA	GAAGAACTGG
7201	TTGTCGATCC	GGAGGAAGAA	CACCATCCAG	CCCATGAGGG	ACCTCCTGAA	GAGGGCCAAC

Figure 8 continued

7261 ATCACTGGGT TCGATTCCGA GACATITATG AATGAGGCCG CCAACAACAT CTCGCAACTG
7321 CCCAATGTCG CCATTGCCAT TTCAAGGAGGC GGCTATCGTG CCCTCATGAA CGGCGCCGGC
7381 TTCGTTGCTG CTGCGGATAA CCGAATTCAA AATACCACGG GCGCAGGTGG TATTGGAGGC
7441 TTGTTGCACT CCAGCACATA TTGTATGTA AAGTGGTTCT TCTTATCTCG TTTTCGAGTG
7501 TCAACTGCGC CAGTTCAAG AGGGCGGCT GTGGACGACC TTGCTGGTGA ACATGTCTTG
7561 GACTCCATGC CCCTTCTTCG TTTCCTCAA TCAAGAACGTC GAGGACCGTG ACCGTAATC
7621 GCTAACGCAA CTCTAGGGCC GGACTTTCTG GTGGTGGCTG GCTTGTGCGC AGTTGTTCT
7681 CCAACAACTT CAGCAGCATT GAGACCCTGC TGAGCGAGAA CAAAGCTGG GACTTTGAGA
7741 ACTCCATCTT TAAAGGGCC AAGGAGGCTG GCCTTAGTAC TGTCAACCGC ATTCACTACT
7801 GGTCCGAAGT GGCAAAAGGA GTTGCCAAGA AGAAGGATGC TGGCTTCGAG ACAAGTATAA
7861 CAGACTACTG GGGCCGAGCA TTGAGTTACC AACTGATCGG AGCCGATATG GGCGGCCCGG
7921 CTTACACCTT CTCCAGCATT GCCCAGACCG ACAACTTCCA GAAGGCCGAA ACGCCGTTCC
7981 CTATTCTGGT AGCTGACGGC CGCGCGCTG GAGACACCAT CATCTCCCTC AATGCTACCA
8041 ACTACGAGTT CAACCCGTT GAGACGGGTA GCTGGGACCC GACCGTCTAT GGCTTGGCGC
8101 CGACCAAGTA CCTCGGGGCC AACTTCAGCA ACGGCGTGT CCCATCGGGA GGCAAGTGC
8161 TTGAGGGTCT CGACCAAGCC GGCTTCGTCA TGGGCACCAAG CAGCACGCTC TTCAACCAGT
8221 TCCTTTGGC CAACATCTCC AGCTACGAGC GTGTTGCCCG ACGTGCTCAT CGAACGCCGTG
8281 ACTTCTGTCC TCAAGGAAT CGGCGCCAAG AGGACGACGT CTCCCCAAATC ATCCCTAAC
8341 CGTTCTGGT CTGGAACAAC CGGACCAACC CCAACGCCGA CACGCTCGAG CTCGACCTGG
8401 TCGACGGCGG CGAAGATCTG CAGAATATTC CGCTCAACCC GCTCACCCAA CCCGTCGCG
8461 CCGTGGACGT CATCTTCGCT GTCGACTCGT CGGCCGACGT GACAACACTGG CCCAATGGCA
8521 CGGCCCTGGC AGCCACCTAC GAGCGCACTT TC GGCTCTAT TTCCAACGGG ACACCTTCC
8581 CCTCGATCCC CGACGACTGG ACGTTTATAA ACCTAGGCC CAACAAACCGC CCCTCTTCT
8641 TCGGCTGCGA TGTTAAGAAC TTACCTTGA ACGCCAACCA AAAGGTTCCC CCCTTAATCG
8701 TCTATGTCCC CAACCGGCC TATACCGCGC TGAGCAACGT GTCCACCTTC GATCCGTAT
8761 ACACCATGTC TCAGCGCAAC GACATCATCG GCAACGGATG GAACTCAGCC ACGCAGGGAA
8821 ACGGCACGCT GGATTCCGGAG TGGCCCACTT GCGTCGCCCTG CGCGGTTATC AGCAGGAGCT
8881 TAGATCGGTT GGGCAGGCCAG ACCGCCAGCCG CGTGCAGAAC TGCTTTGAG AGGTATTGCT
8941 GGAATGGCAC AGTGAACCTCA AAAGATACAG GGGTTTACAT GCCTGAGTTC AAGATTGCGG
9001 ATGCCATGC CCTGGACTCG GGTGCTGTTG CTATCGGAAA GATGGTGAAT GTCTGGTGT
9061 CGGTTGTGT GGGAGTTGTG GCGGCTACTT TGTTGTTGTA GGGGTTAGGGG AGACGTGATG
9121 ATATTCCAGT CTGATGAAGT TGAGACTGGA CTGGAGATCG CCAAGGATGC GGAGGGAAAG
9181 GAATGCGTGG TGTTAATGTC ATGATGGATG AAGGGTCATG GATCATGGAA CGACGGGGCG
9241 GGGATATTGG ATGATGGATA TACCACACTG CATGCATGCT CTATTGATAA TATGCTTTGG
9301 CATTACGTT TAACAATCAA TTGCTCCATC CTGATGTTCT ATCTTTGAC ACTGGATTGA
9361 TACTACTCCT GTGCTTCCC TCTTGAAGTT GGAAGGACTT GAGGTTGGAA GGACTTGAGG
9421 TTGTTGFTC TGAGGGAGGT TATCGAAGTA TCATCTGTGC TGATGCCGAT CGATAGACTG
9481 CCCTCTTCTT CGAGGCAACG AACGGTCGGA TGAGCCTCTA ATCATGATGC TCAGTGCCAC
9541 AAAAAGGCTC CAGCACAGCT GCCCACACCT TTTTGCCTC GTCGCTCCCT CCTTTTTTTC
9601 CCCCCCTTTC TTCCCTTCCA TCTCATCCCG TACCAGAGTG CCCACCGGGT ATATATAATTA
9661 CCTCTTGGC CGTTCTCCCT TGACCAATAA ATCGCTTGGT CGAGTGGCGT AACCGTTTAC
9721 CGTCTACACT TATCACTCAA ACCAAACCAA ACCATCGAAG AAGTTACCTA TCGGTTGAG
9781 GGAACGGTGA FGTTCTTACG TTCAAGTTAA CCCAAAGAGC GTTCCACATC GTTGAACCGT
9841 CTCCTCCAGT TCTTGGATCT GTTAACTTC CGCAGCGACT GAAGAAGTAA TCACTTTTT
9901 TTTTTTTGGT TCCAAAAAAA AAAAAAAA TTAC

Figure 9 Construction of the components of the sequence diversification cross: Parent (variant 1) and Parent (variant 2). For convenience, plasmid sequences are shown as linear. The cross hatched region in the chromosome in the foreign DNA is dispensable. Stippled sequences in the plasmid indicate the multiple cloning site for inserting foreign DNA. Crossovers in region 1 and region 2 insert the foreign sequence to be diversified into chromosome 1 of *Neurospora crassa* adjacent to the recombination hotspot *cog*. Parent (variant 2) containing a version of the foreign sequence with multiple differences from that in parent (variant 1) is similarly constructed. Parent (variant 1) and parent (variant 2) are crossed and conversion events (stippled arrow) initiated (X) at *cog^L* recombine the sequence differences in variant 1 and variant 2 to form new combinations. Sequences are identical except for those that distinguish variant 1 and variant 2. *rec-2* on linkage group V permits *cog^L* to be active. For simplicity, genes not directly related to the diversification are omitted. See text for further details.

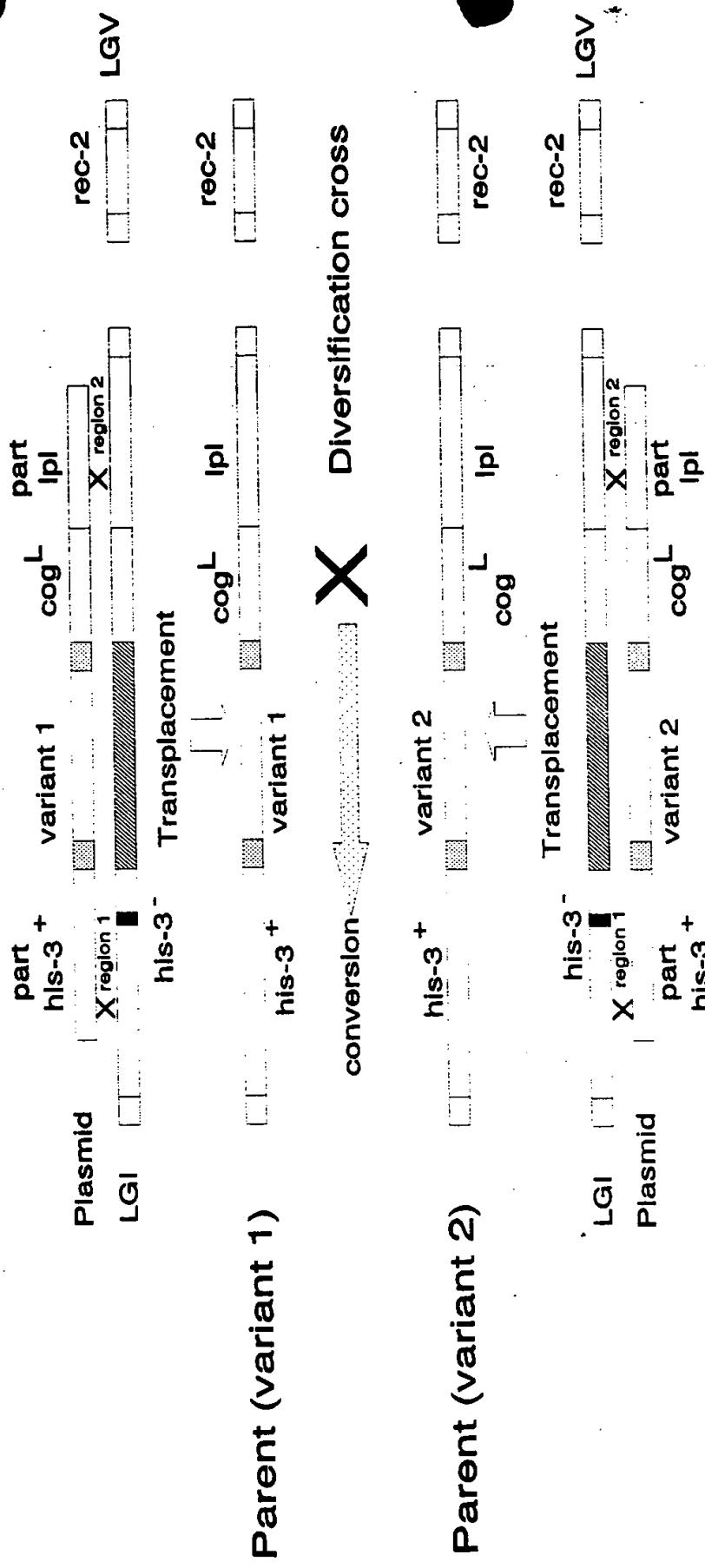


Figure 10 Construction of parent (variant 1) and parent (variant 2) enabling selection of progeny that have experienced conversion in the foreign DNA. Complementing pairs of *his-3* alleles are used to obtain parent (variant 1) and a different pair of complementing *his-3* alleles are used to obtain parent (variant 2) as explained in the text. Parent (variant 1) and parent (variant 2) are crossed and his^r recombinants are selected. These must all have experienced conversion events affecting the foreign DNA since the events begin at *cog^L*. The *his-3* alleles in parent (variant 1) and parent (variant 2) are non complementing to ensure that selection yields recombinants and not aneuploid progeny having two copies of all or part of linkage group I.

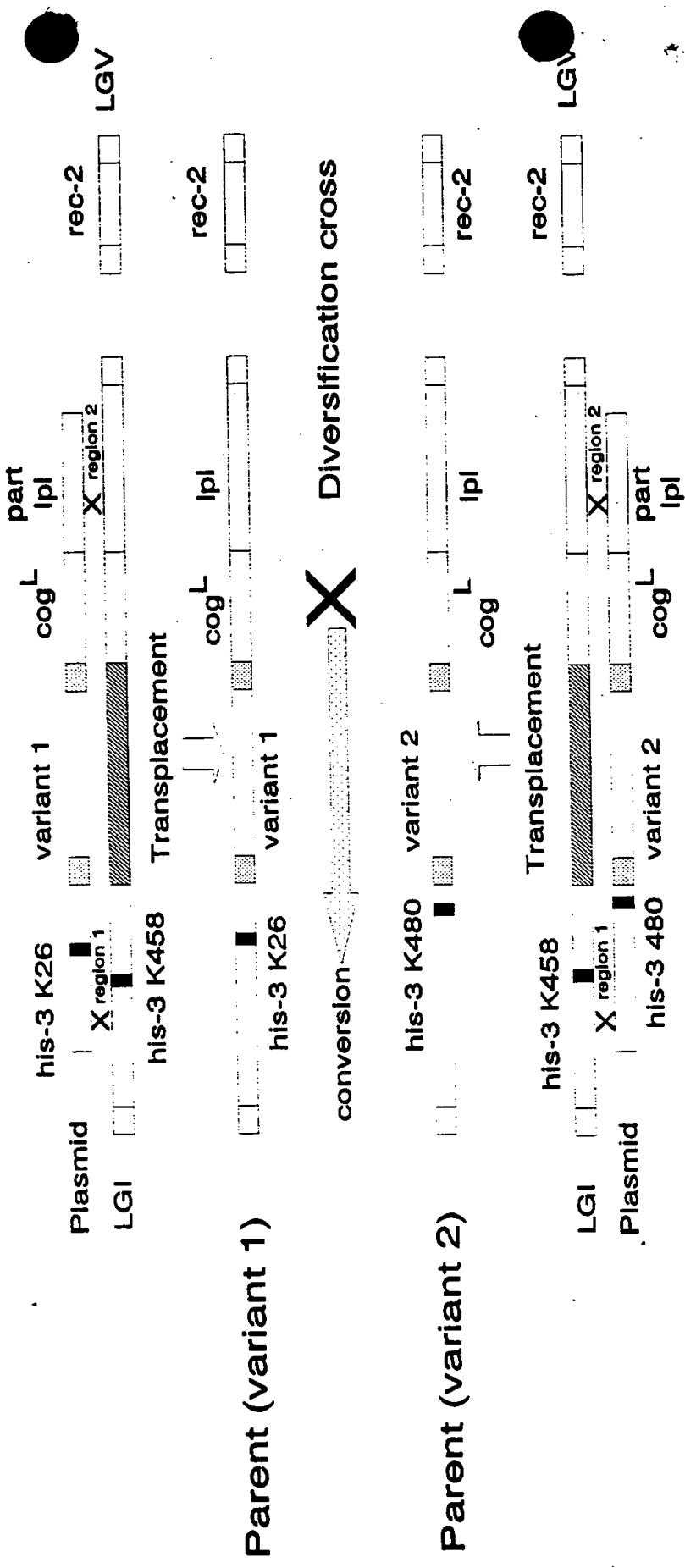
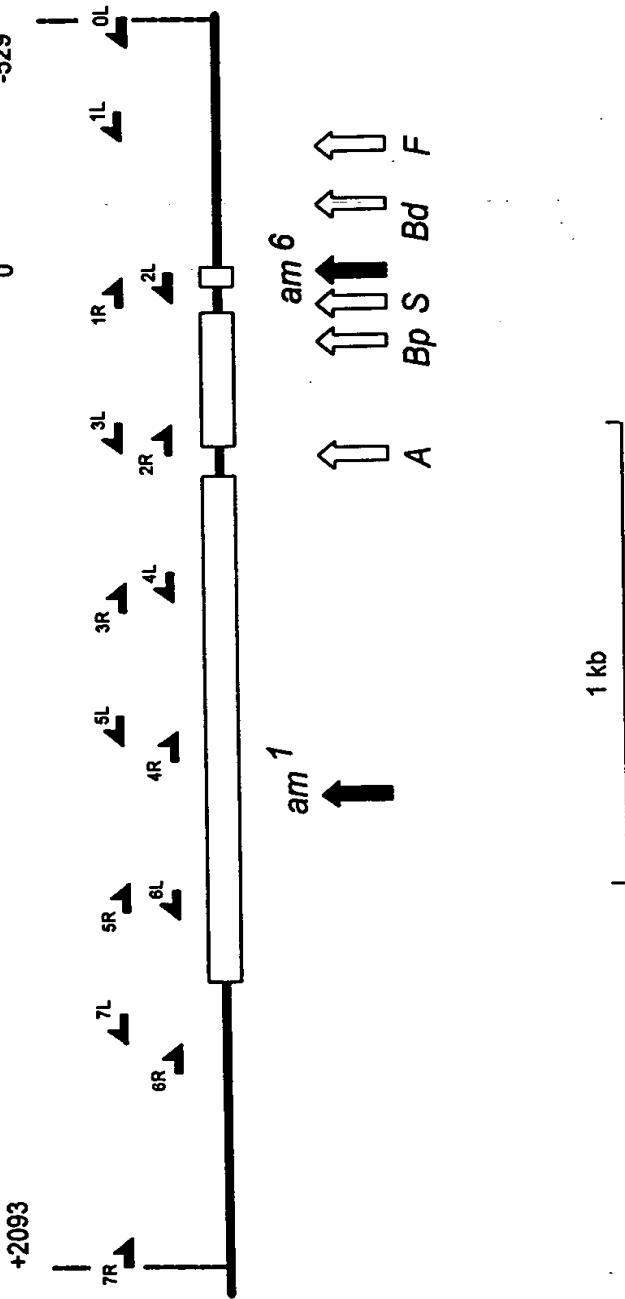


Fig 11



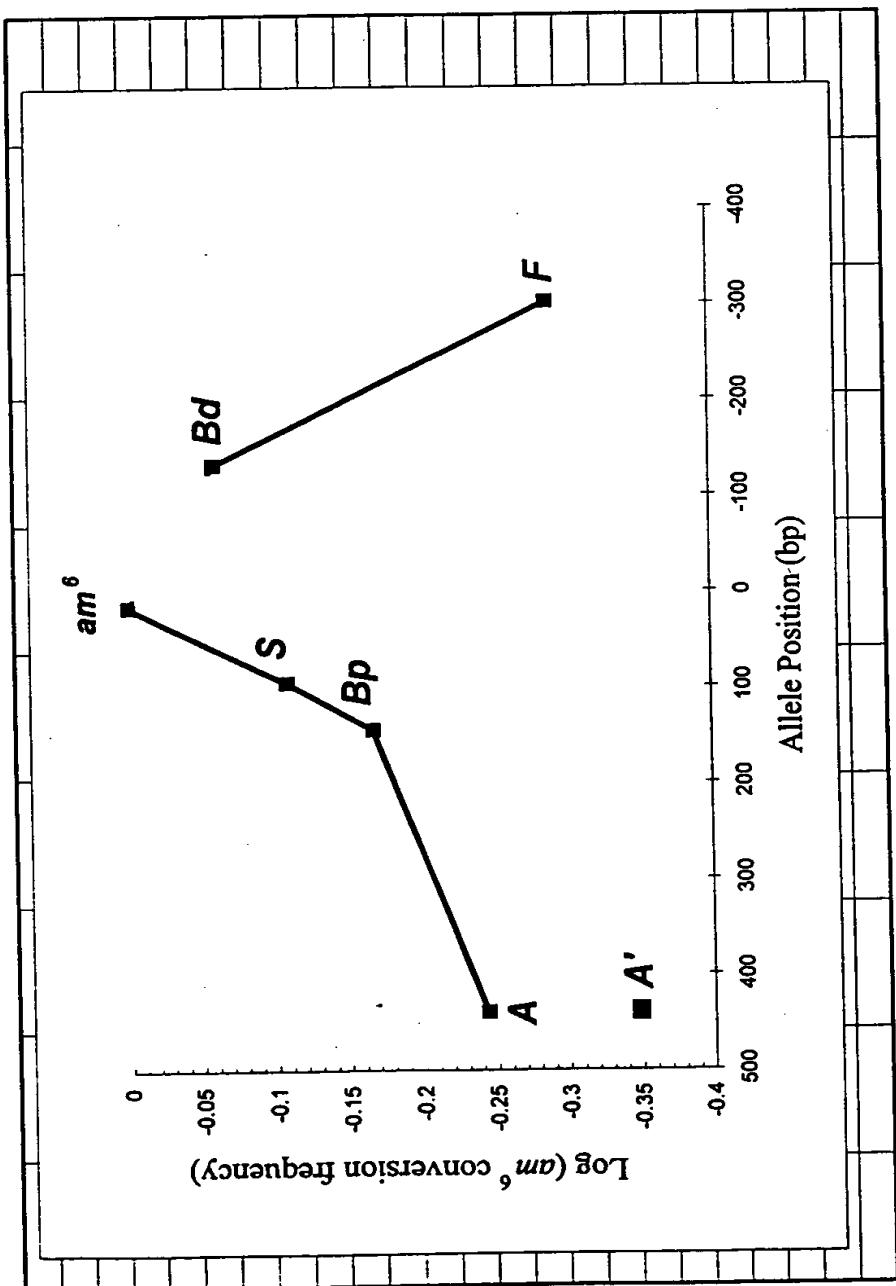


Fig 12

Fig 13

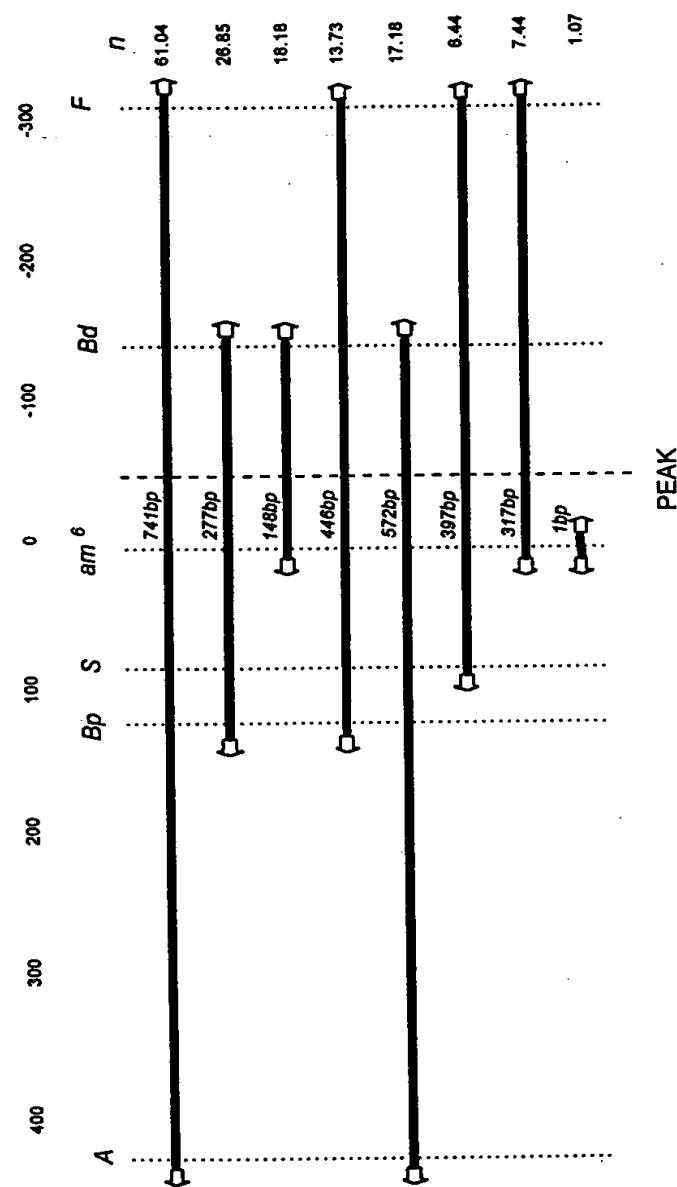
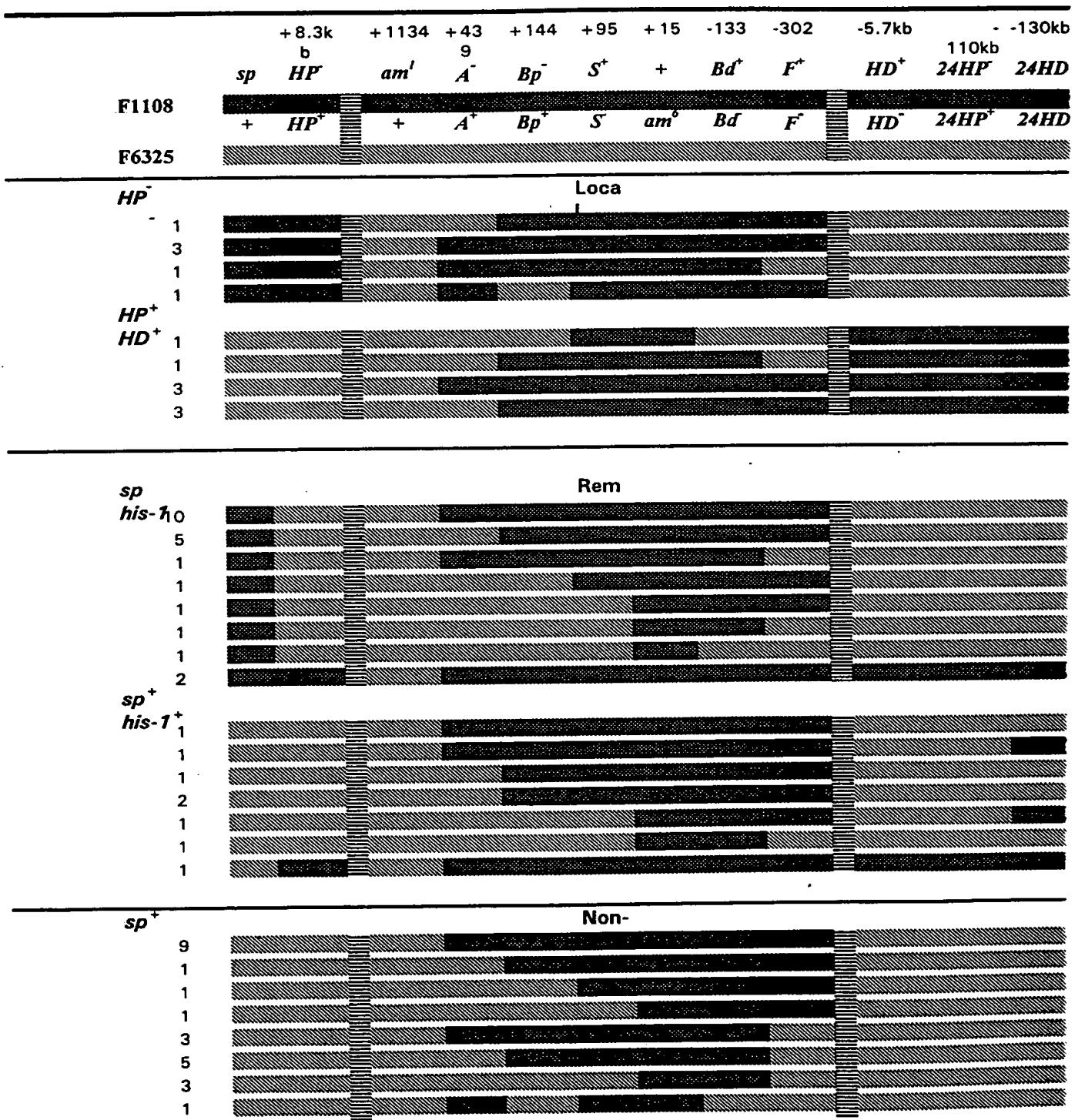


Fig 14 (2 pages)



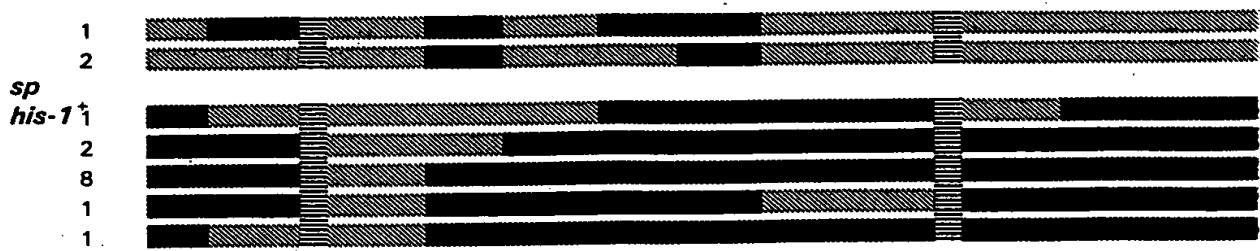


Fig. 14 (continued)

FIG. 1A

Methods for the diversification of DNA sequences and testing for superior variants

existing protocols: Number of transfections needed to generate 1024 new variants: **1024**

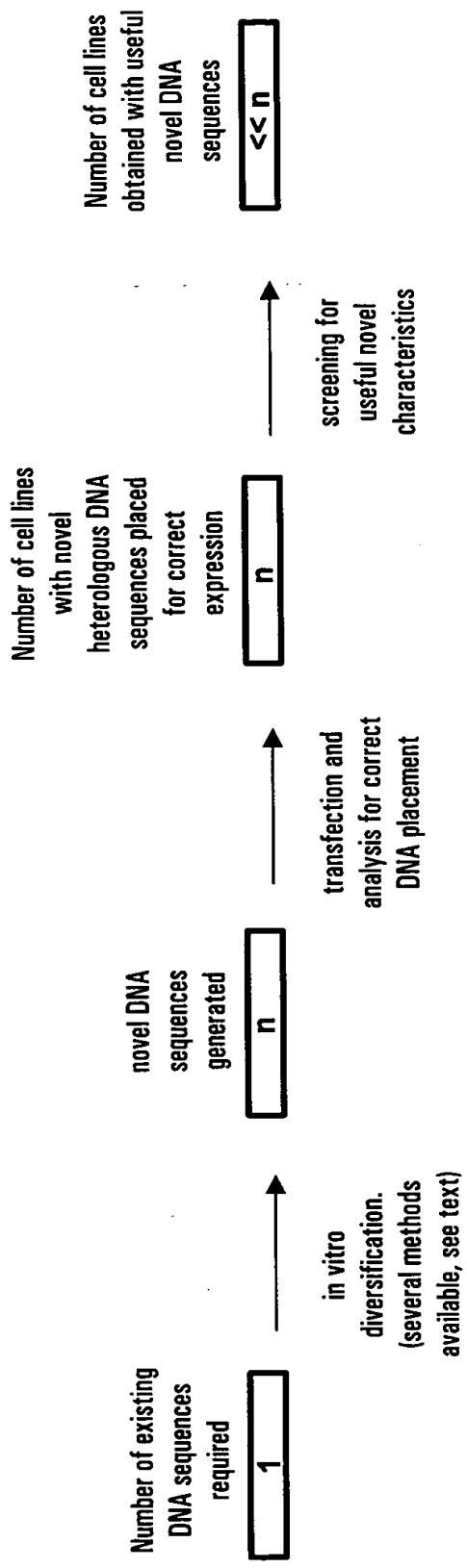


FIG. 1B

A protocol enabled by the present invention: **Number of transsections needed to generate 1024 new variants: 2**

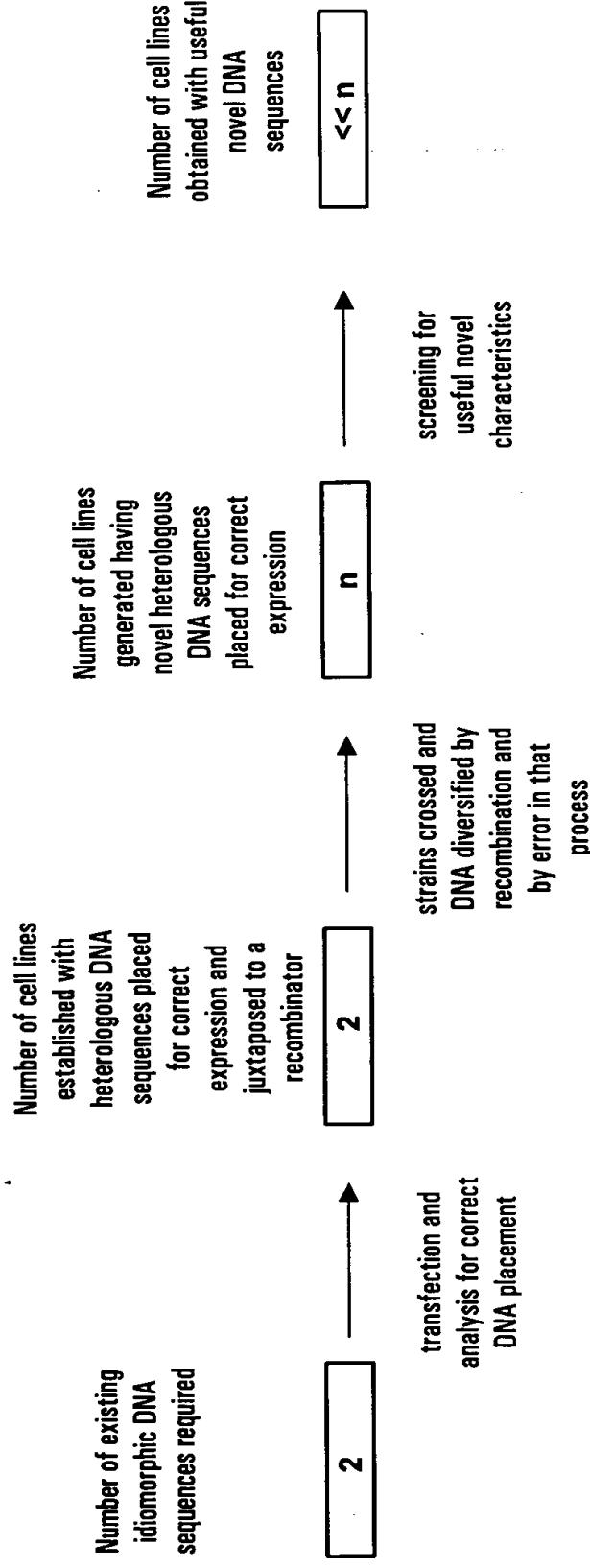


FIG. 2A

Methods for the diversification of DNA sequences coding subunits of heteropolymeric proteins and testing for superior variants.

The example given for immunoglobulins is for illustrative purposes only and is not intended to limit application of the present invention to this specific heteromeric protein. H = heavy chain genes, L = light chain genes

Existing protocol: Number of transfections needed to generate 1024 new combinations: 2048

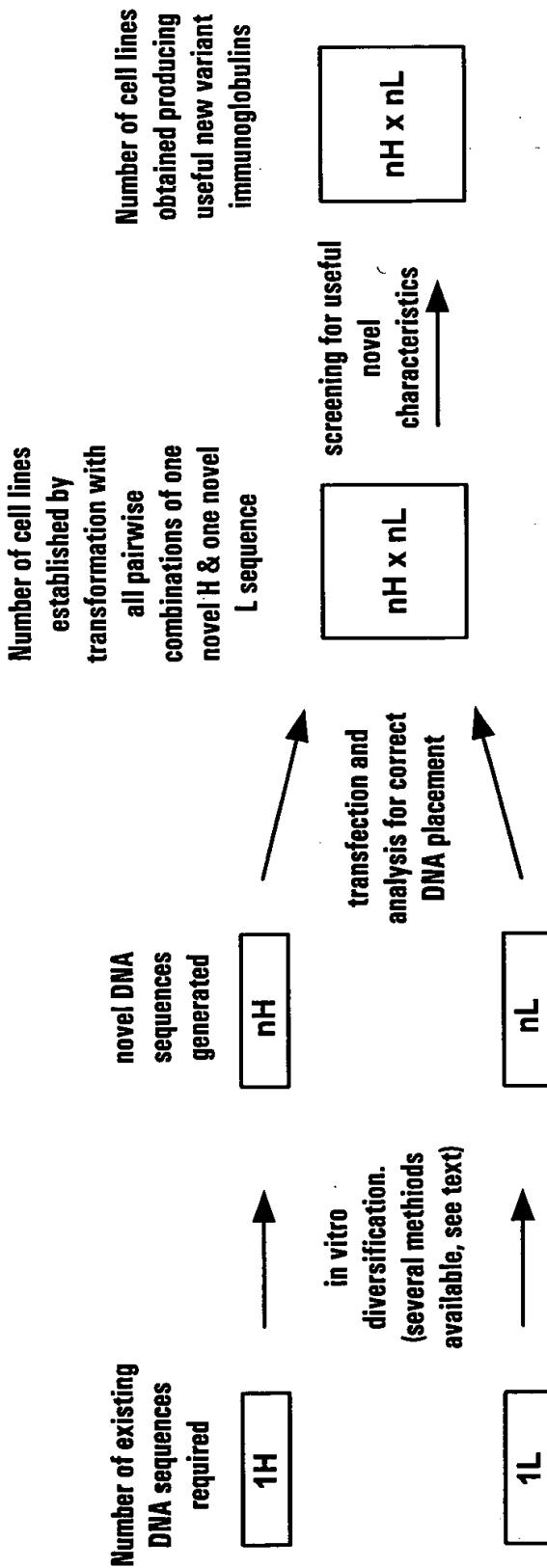


FIG. 2B

**Existing protocol using the heterokaryon technology of US Patent Serial No. 5,643,745
Number of transfections needed to generate 1024 new combinations: 64**

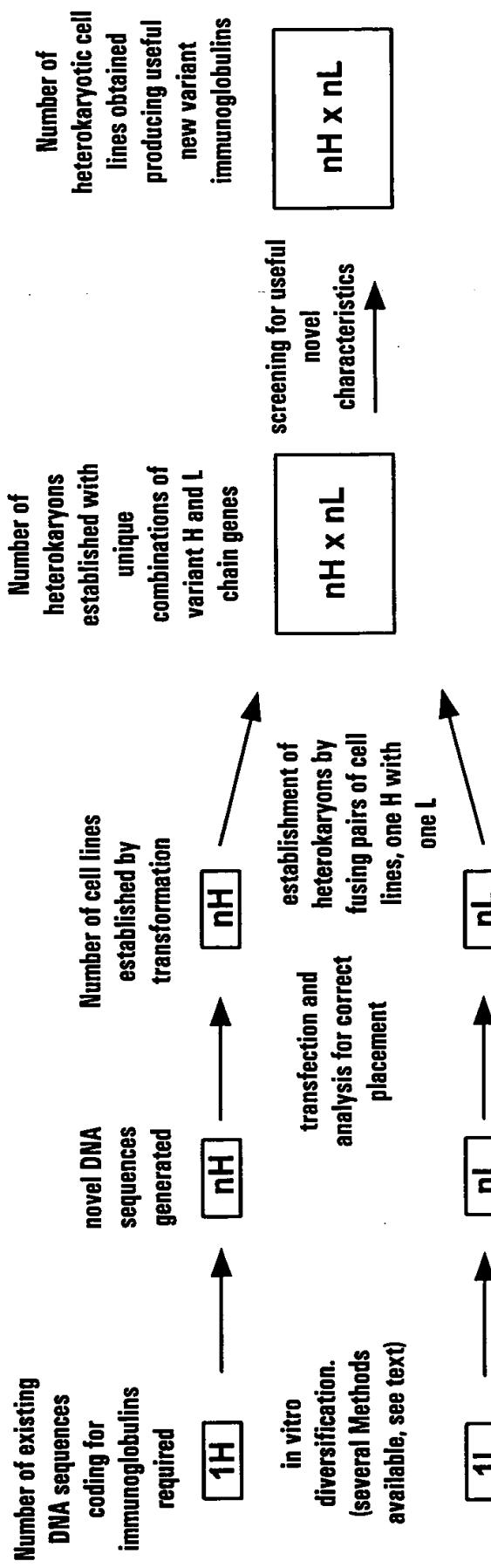


FIG. 2C

A protocol enabled by the present invention

Number of transfections needed to generate 1024 new combinations: 4

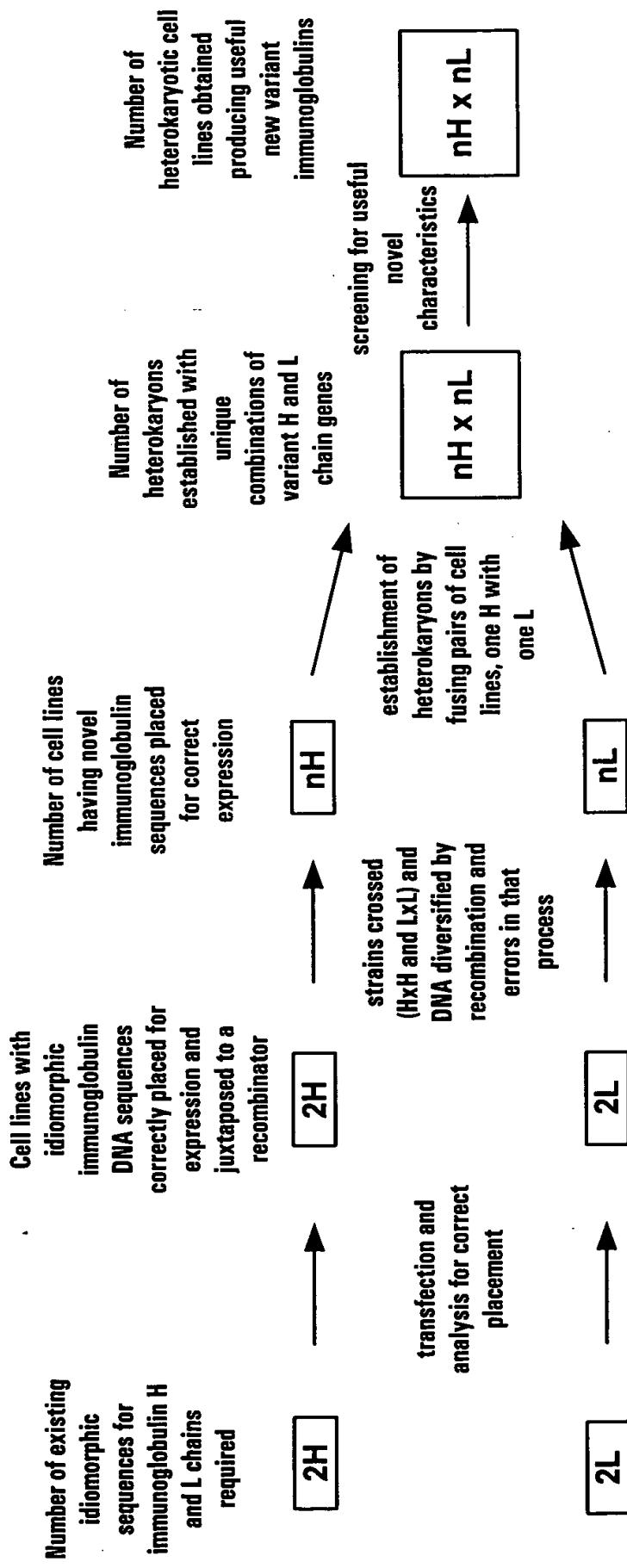


Fig. 3

The modified double strand break repair model for meiotic recombination. After H Sunet *et al* Cell 64: 1155-1161, 1991

(a) A double strand break (DSB) is made in one DNA duplex. (b) A long 3' overhanging single strand tail is generated either side of the break by resection. (c) One 3' end invades a homologous duplex forming a D loop. (d) the D loop is enlarged by repair synthesis and anneals to the second 3' end (e) Repair synthesis occurs at the second 3' end and two intermolecular junctions (Holliday junctions) are formed. Resolution of the junctions by cutting inner and outer strands can give rise to non-crossover (f) and crossover (g) chromosomes. If there are base mismatches in the heteroduplex regions (duplex molecules with thick and thin lines) there will be gene conversion. If mismatch repair does not occur there will be post meiotic segregation of new sequence combinations.

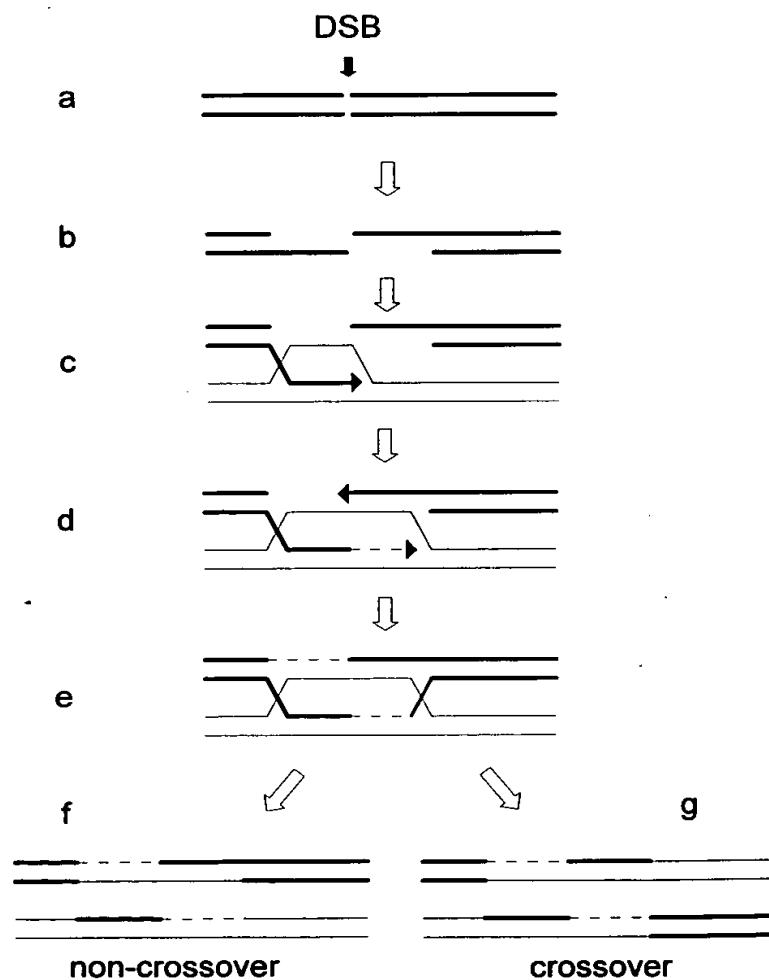


FIG. 4

Life cycle of *Neurospora crassa* after JRS Fincham (Genetics, Wright 1983). Microconidia having one nucleus are not shown but can be generated as described in the text. Perithecia and protoperithecia are shown in section.

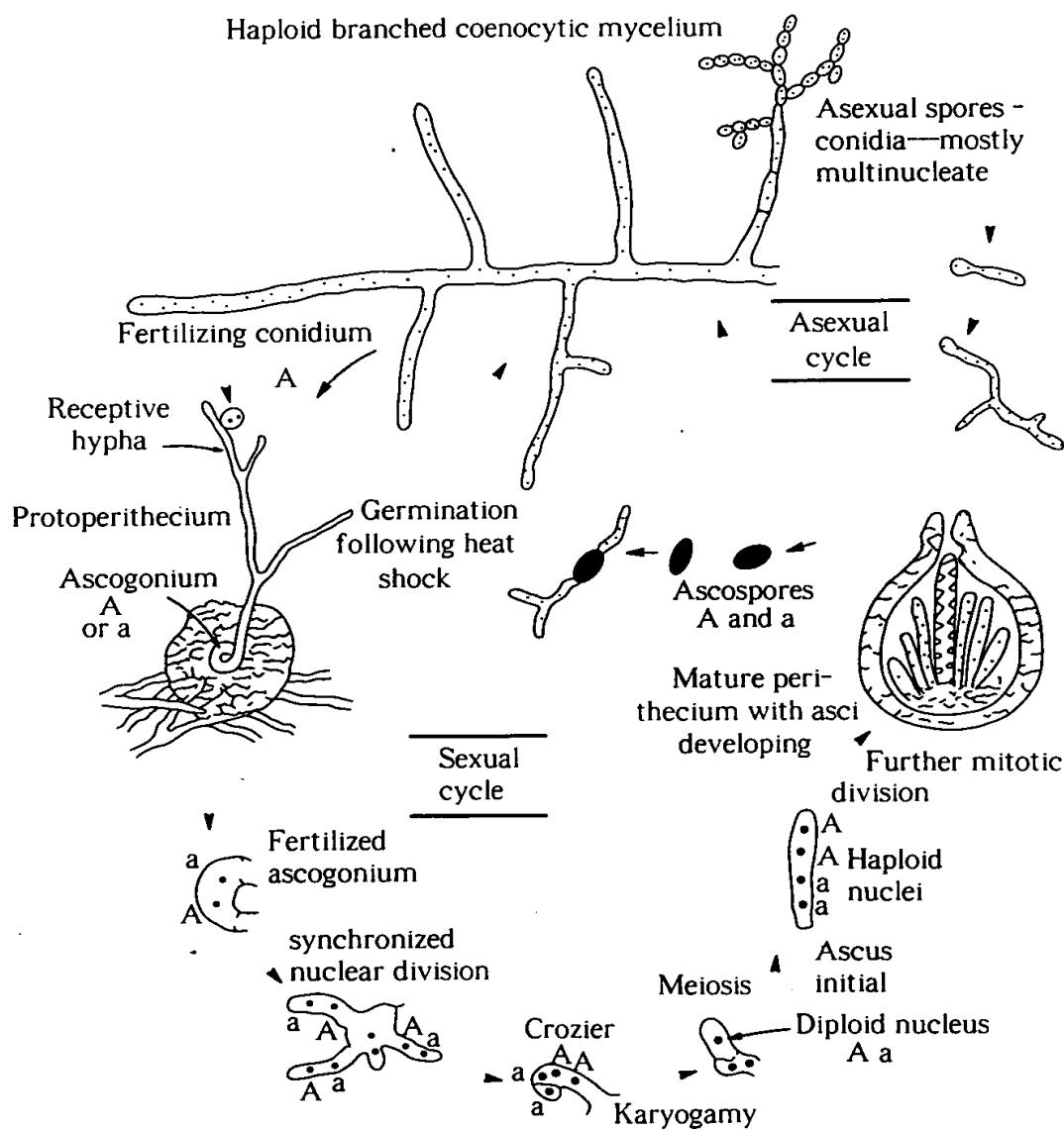


FIG. 5 Map of the *his-3*, *cog*, *lpl* region of Linkage Group I of *Neurospora crassa*. Vertical bars, triangles and hairpins show the location of sequence differences that distinguish the St Lawrence and Lindegren wild type strains. The corresponding full DNA sequences are given in Fig. 7 and Fig. 8. Vertical slashes indicate one to seven base substitutions per 10 base pairs. Triangles indicate short sequence insertions and the hairpin a 101 base pair inverted repeat transposon fragment present in St Lawrence.

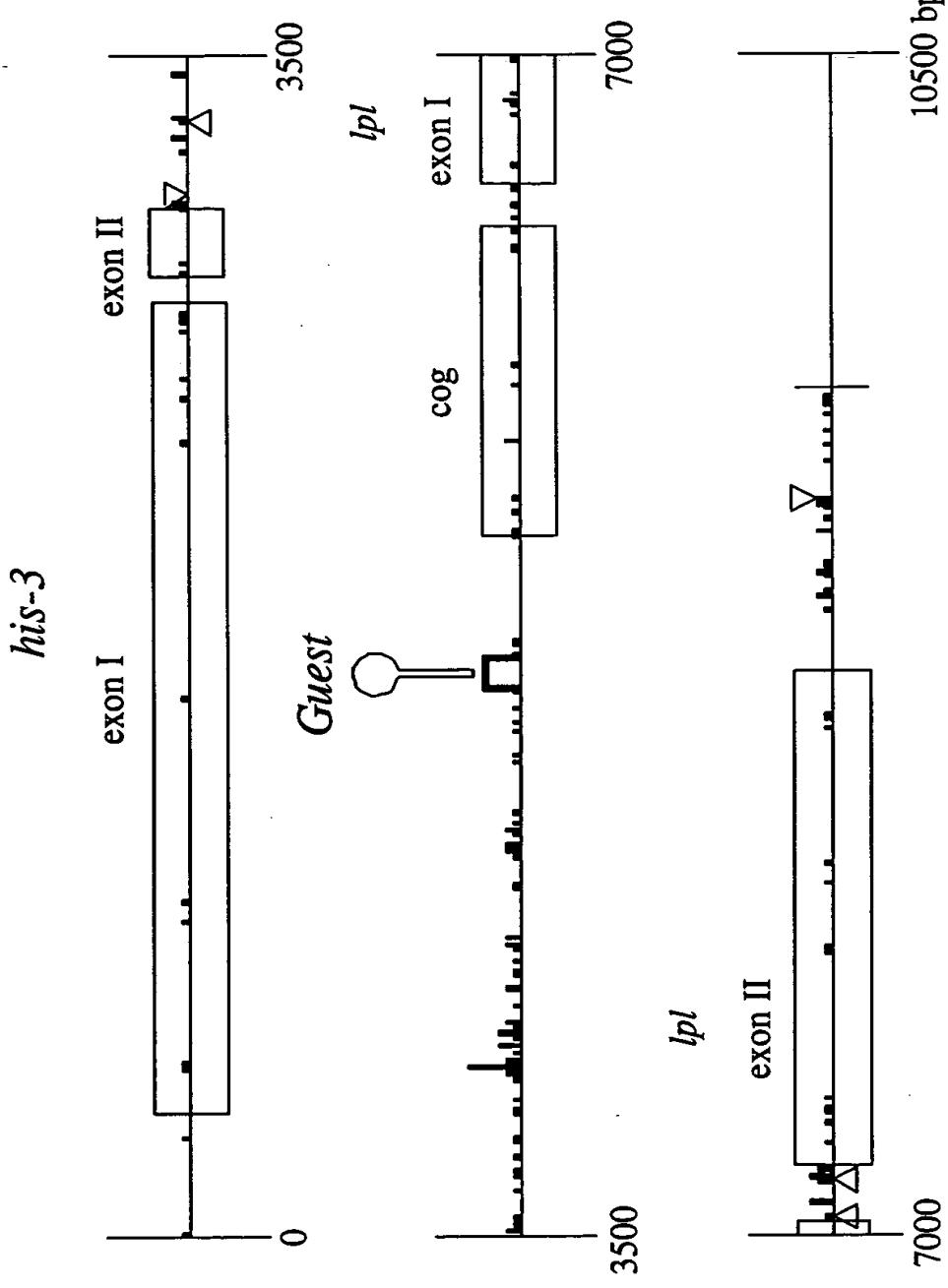


FIG. 6

Discontinuity in the parental origin of DNA sequences in progeny from crosses between pairs of *his-3* alleles. In most cases this reflects discontinuity of conversion tracts, in some cases crossovers near the ends of conversion tracts.

Markers are specific DNA sequence differences that distinguish the parents. These were all E (Emerson wild type origin) or all L (Lindgren wild type origin) in the parental strains. Recombinants carry both E and L markers. Marker position is given in base pairs from the first base of the first codon of the ***his-3*** gene. Each line of the table shows the parental origin of the markers inherited by one of the progeny.

marker location	P	H	P1	K1201	K504	L3	R1	K26	K874	R4	C4	C5	C6	C6/7	C7	C8	C1	C2	C3	C9	D
-60000	-384	115	179	563	1232	1502	1717	2318	3436	3705	4000	4304	4667	4821	5232	5495	6153	6507	cog	his-3	

FIG. 7

Nucleotide sequence of the *his-3 cog^L lpl* region of linkage group I in the Lindegren wild type strain of *Neurospora crassa*. This differs from that in the StLawrence strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the high frequency recombinator *cog^L* which is active providing the cross in which meiosis occurs is homozygous *rec-2*.

1 GATCGCAACT GGAGATCACT CGCACCGTGC CGCAGAACAA GGGCGACGAG CCTCAGGGCA
61 GTTTAGCTG CCGTAACAGC ACAGACCATA GCTTATTTC ACCTGGCGG GCGGGCAGC
121 GCGGCACTGA CATCGGCAAG GCGGCATCAA GCAACCCCTC TGGTGCTTGC CAGCTGCCGG
181 CCAACGTCAG CGGTACAAGG AGAAATCTGG AAGGAAAGAC TTCTGGCACC GACAGGATGG
241 CACCGGGAA AAGTTCCAA TGCAATGAGAT GAGGGGCATT TGCATTGCCT CCCGTACAC
301 TGCCCGGAA CCCCAACCCC ACCATAGCGT CTGTCGATAC ATGGAGCGCG AAGTCGAGAA
361 ACCTGTAATT CCTGGTAACT TTCAGGTACA CAGTACGTAC TGATCCTGGT ATCAAACCTT
421 GCCTGCCGAG TTTTCGACGG AAAAGAGGTGT GAATTGTGAA AGAGTCATAC CAAATCACCC
481 GATTTCTATA AAGCCCCAGT CTTTCTGTA CATAAGCGAC ACTCGAAGCG GGCCTCATCT
541 TCATAGCCTG ATAGCTTGT AATACTCCATC CTCGTATCTC ACTTGACCTT GAGTTCAACC
601 CCACGTCAGA CTTCACCCGA CACATCGACG GATTGGGAA CAGCACAATA CCTGAAAAGC
661 GAGAAAACCA AACAGAGGAA AACACCATGG AGACAACACT TCCCCTCCCC TTCCCTCGTCG
721 GTGTCAGTGT TCCTCCCGA CTGAATGACA TCAAGGAGGG CCTCAGGCCGG GAGGAAGTCT
781 CGTGTCTTGG CTGCGTCTTC TTGAGGTCA AGCCCAAGAC CCTTGAGAAA ATCGTGCAT
841 TCCTCAAGCG TCACAATGTC GAATTGAGC CCTACTTCGA TGTAAACAGCC CTCGAGTCTA
901 TCGATGATAT TATCACTTT CTGGACGCCG GCGCCCGAA GGTGTTGTC AAGACCGAGC
961 AGTTGGCCGA CCTCTCCGCA TATGGCTCCC GCGTTGCCCG CATTGTCACT GGAAGCAGCG
1021 CTGCTTTGCT TTCCCTCCGCC ACCGAGAGCG GCCTTTGCT CTCCGGCTTC GATCAGACTG
1081 CCTCCGAGGC TGACACAGTT CTGGAGGAGG CCAGAGACAA GAAAATTACC CCCTTCTTCA
1141 TCAAGCCCCGT TCCTGGGCC GATCTCGAAC AGTTCATCCA GGTGCGCCGCC AAGGCTAACG
1201 CCATCCCCAT CCTGCCATCC ACTGGCTTGA CAACAAAGAA GGACGAGGCC GGAAAGCTTG
1261 CCATCTCCAC CATCCTCTCG AGCGTCTGGA AGTCTGACCG TCCCGATGGT CTGCTCCCCA
1321 CCGTTGTCGT TGATGACAC GACACTGCTC TGGGTCTGGT CTACAGCAGT GCCGAGAGTG
1381 TGAACGAGGC CCTCAGGACA CAGACTGGTG TCTATCAGAG CCGGAAGCGC GGTCTCTGGT
1441 ACAAGGGTGC TACTTCCGGA GACACTCAGG AGCTCGTCCG CATCTCGCTT GACTGCGATA
1501 ACGATGCTCT CAAGTTGTC GTGAAGCAGA AGGGTCGTTT CTGCCACCTC GATCAGTCCG
1561 GCTGCTTTGG TCAGCTCAA GGCCTCCCA AGCTCGAGCA GACTTTGATT TCGAGGAAAC
1621 AGTCTGCCCG CGAGGGCTCC TACACTGCC TCTCTCTCTC CGATGAGAAAG CTAGTCCGGG
1681 CCAAGATCAT GGAGGAGGCT GAGGAGCTCT GCACCGCTCA GACCCCCCAG GAAATCGCCT
1741 TTGAGGCTGC CGATCTCTTC TACTTGCTC TTACCAAGGG CGTTGCTGCC GGC GTTACTC
1801 TTGCCGATAT CGAAAGGAGC CTTGACGCCA AGAGCTGGAA GGTCAAGCGC AGGACTGGAG
1861 ATGCTAAGGG TAAGTGGGCT GAGAAGGAGG GCATCAAGCC TGCGGCGTCC GCTCCCGCTG
1921 CCACTTCGGC CCCTGTCAACC AAGGGAGGCCG CCCAGGAGAC CACCCCTGAG AAGATCACCA
1981 TGAGACGTTT CGACGCCCTC AAGGTCTCTA CCGAGGAGCT CGATGCTGCT CTCAAGCGTC
2041 CTGCGAAAAA GTCGTCCGAT GCCATCTACA AGATCATTGT CCCCCATCATC GAGGACGTCC
2101 GCAAGAACGCG CGACAAGGCT GTTCTGTCGT ACACTCACAA GTTCGAGAAAG GCTACCTCTC
2161 TTACTAGCCC CGTCCTGAAG GCGCCCTTCC CCAAGGAGCT TATGCAGCTC CCTGAGGAGA
2221 CCATTGCTGC CATCGACGT TCCCTCGAGA ACATCCGAA GTTCCACGCC GCCCCAGAAGG
2281 AGGAGAAGCC CCTCCAGGTC GAGACCATGC CCGGTGTTGT CTGCAAGCCGT TTCTCTCGTC
2341 CCATCGAGGC CGTCCGCTGC TACATCCCCG GCGGTACCGC CGTTCTCCCC AGCACTGCC
2401 TTATGCTGGG TGTTCCGCC ATGGTCGCCG GCTGCAACAA GATTGTGTT GCCTCTCCTC
2461 CCCCGGCCGA CGGAACCATC ACTCCCGAGA TTGTCCACGT CGCTCACAAG GTTGGGCCG
2521 AGTCCATCGT GCTTGCCTGGC GGTGCCCAGG CCGTAGCTGC CATGGCCTAC GGCACCGAGA
2581 GCATCACCAA GGTCGACAAG ATTCTCGGCC CCGGTAACCA GTTCGTCACT GCTGCCAAGA
2641 TGTTGCTCAG CAACGACACC AACGCTGCCG TTGGGATTGA CATGCCGCT GGGCCGTCCG
2701 AGGTGCTGGT CATCGCTGAC AAGGACGCCA ACCCCGCGTT CGTTGCCTCG GATCTCTGT
2761 CCCAGGCTGA GCACGGCGTT GACAGTCAGG TCATCCTGAT CGCTATTAAC CTCGACGAGG

FIG. 7 continued

2821 AGCATCTTCA GGCTATTGAG GACGAGGTTG ACCGTCAGGC TATGGAGCTT CCTCGCGTCC
2881 AGATTGTCCC TGGCTCCATC GCCCACTCGA TCACCGTGCA GGTCAAGAGCC GTCGAGGAGG
2941 CCATGGAGCT CAGCAACAAG TAGCCTCCTG AGCACTTGAT CCTCCAGATC AAGGAGGCCG
3001 AGAAAGCTGT CGATCTTGTG ATGAACGCTG GTAGTGTCTT CATTGGCGCT TGGACTCCTG
3061 AGTCCGTTGG CGATTACTCT GCTGGTGTAA ACCACTCGCT GCGTAAGTTA CATATCATAA
3121 ATAGCCCCGC TTCACAGATT CTTCTGCTAA CGTCAAGACA CATAGCTACC TATGGTTTG
3181 GCAAGCAGTA CTCTGGCGTC AATCTCGCT CGTTCGTAA GCACATTACC AGCTCCAAC
3241 TGACTGCCGA GGGTCTCAAA AACGTCGGCC AGGCTGTCT GCAGTTGGCT AAGGTTGAGG
3301 AGCTCGAGGC TCACAGAAGG GCGGTCAAGCA TCCGTTGAGA GCACATGAGC AAGAGCAACT
3361 AGACGGAAAT TCTTTTCGA AGTTGCAAAA AAAACAAGAA CAAAAGGATG TAGTGGGTTG
3421 ATGTATATCT GGGTCATTG GGGCACATAG AGTAATGATA ACGAGTTTG GACATTGTAC
3481 TGTCTGTAC AGGCTGAAGA TCAGTACATG AATCTGTTGG TAAGTGTAGA GACCAAACG
3541 TCCCTTGAGT TTTTCTCCCT GTTCCAGAGA GGTGCTCGTC CCTGGGTGTT TATTTTCATT
3601 ATTACATCAA CCTTTTATT TATTTTATT TTTATTTAC TTTTTTTCC TTTTTTCAG
3661 ATCATGCGTA CATGAACGGG GGAAGCACAG ACGATCGAAA CGTGGATGTC ACAATGTCGC
3721 TGCAGTGTAT CGCATTGCA TGAAGCGCCC ATCTCAATAT ACTTGAGTC TTGCGCGTTG
3781 CACGTGAAC TCCCCAACAA CCGAATAAAA GACGGCAGAA AATGAAGATA AAAAAAAAC
3841 ATAATAAAA TCGGAGGGAG TGTGGAAAT GTTCTTTT AGCATTAGA CCCCATAGCC
3901 GTGCACGCC GGGTACAGAC AGGTTCATCG ATGTTGACAT TGACTGGGAC ACCAGGTCTA
3961 TCTATTCAT CTCCTGCTCT CTACCATACA TCGGGACATC GGACATCTCG CTGTACCCCC
4021 CACACCCACA AAGTCTTATA AAAGCGCCAC ACCCGAGGAG GTTCGGTCGG CCCCACGAAC
4081 TCCGTGCCTC CCTGCTGTG TACAGGGACC GAACCGCTGGA GAAGCTTAGT TTCTGACAT
4141 CCGGCCTTACCG CGAGCAGGAA AAGGGACAGC TCATAGGCGA GGAGGGATTT GAAGATGGGG
4201 ACATTTGGA TGATTGAGA GGAGGAACTA GGTACTGTAT CATGATAGTT CGGGGCAGCA
4261 TCTTGGCTGG GACATTGTTA ATACCTCGAT ATGATGAAGT GGGAGGGAGT TTTTCATGT
4321 CTTGCCAAG TCCCACATAAT CTTTTTTT TTTGTACCA ACACCAAGA TTGGAGAAT
4381 AGTGTAAAGGA TTCGCATTCA CAAGTGGAAAG TCTGAGGATC TTTTATATC TTTGCTTCC
4441 GCGGACTGTT AACGATCCTA CAGCGAGCGA GCGAGCGGTG GGATGCGCTG ATCTGATAGG
4501 TCGAATATAC GGCGCTTTC TCCGGTCGTG TAGTGTAAAGC TCTGTCGGCA TAGTAGTACA
4561 CTAAAAAAAC CCTTGATTT CATGATCTGC TTGCTATTCA TTCCGAGTTA TTTCAGTGGT
4621 CACATTCGA GATTCACAGC CATCCATCCA TATGGAAAAA TCCATTCCA TGCTTCCTCC
4681 CCCCCACTAT GTATGTGACC ACACGCTGCT GTCAGAATGC CAACGGTCTC AGGTACCC
4741 GTCCGACTGT TTGGCATGGA GTTACATACA CTACTAGTGT AGCCCCGGGC CAAGCTACCC
4801 CGTCAAATCT ATACATATCT ATAATGGGTT TCAGGTGTT CGTCGCTGT CAATCAAGTT
4861 TGAAACATCA CTGGGCCGT TGGACGGTGT ATTAGACCAT TGGCTCCCTC AGCTGGCGGC
4921 TGGCGGTTG GGTGGCAAT AACGGGACTG GACTTGAGAG GGACGAGGAG AGTCGGTTGG
4981 CTGCCTACAC TACACTACAA GCGTCCAC CTAACCGACG AGTCCCCTT TCCATTGTG
5041 TGCCCTAACC ATCATCTAGG GATGTCAGGG TTTGCCGGA TCAGGGTATG TTTGGTGCAC
5101 TGTGTCATG TCTGATTGGG TACATATCAT GGTAGGTGTC TCGAGAACAG TAGAGTACTC
5161 GGGCTAGCG TTTGGATGAT TACGCGAGAT ATGAGTTGTA GGCGCCATG CAGTTGCTTG
5221 CCCATAAGCA GAAGTTGCTT TGGGATATAT TTCTGCTT TCAAAGGTCA CGAGGTCTG
5281 GGACGAGCGG CATGCCATC CAAAGGGTG AACATGAGAA ACCGGAATGG CCTTGCGTT
5341 GAAATACAAA AAGTCAAGAA TAAAATCGCT TGAGGATAGG GACGTGGAAG CAAGCAAATA
5401 TGTAAGGGG GGTACTGCTA TGTAGGTGCT CAGCAAACG CCAATTCTT GGCCCCAAG
5461 CAGCAGTTG CTGTCAGTGC TGCTCGTGC AGCCTTGGTA GTGGAACCTA AACTGCTAAC
5521 ACAGCGCAAG TGCAGTGTAA AAGATATTGT GGGAGGATCT GTATGGATGG ATGAGATTAC
5581 TGCTTGGTGT TGGTTGCGAG GCACTGCGC TGTTAGGCTT TGCTGTGCC CGTTGACGA
5641 AGAAATACCG GGAACATATAA ATTGGATACC TAGACTTACT GCCTATGGGA GGTATCTACC
5701 GACGTAGCCG ACGGATTCTA GCAACATCCC GACTTGCTT GTAGTGTACT ATGATAGCAG
5761 CACAGTGGGG TGTGCTCCT TGTGAGCATG GGCTTTTT TTTTTTTCC CCCTCCCTA
5821 GGGCGTTGAC TGGACTTGCT CTATCGTTCC CAAGGTAGGT GCCCGTCATC GATTTCCCA
5881 AGCCGTCTCC CGCCAGATTG TCGTCATAGT GTCATGATGA CCTCGGTCGC TGGGGCTGCG
5941 TGTTACGGG GAGCTGGGAC CGCTAGGCCT CAGTGGTTGT GCCATTCAAGC GTGGGTGTGT
6001 GGAGTAGCGG TAGAGGCGCT TGGAAAGTTGT GCTAGCGAA ACCCTGGAAT ATCTTGTACC

FIG. 7 continued

6061 CTTCGATTCC TTCTCGGGCT GCCCATGTGC TGAGGTGATG CGGGGGATCT GGCGCCAATC
6121 ATCCATTGAG GTTCCCGCAG CTTCCCGBTG CGCGCGCGG GCGCAGTTGC TCACAGGACA
6181 CACCTAGACG CAGGGGCACA GGGGCACCGT TTGGTGTGCA ACTGGGTACC TGGTAGCTGT
6241 AGCAAGCACT CCACCGTCTG TGCAATCCCC CAATCCACGG CAGGAACCTA GCACCGCCGC
6301 GGCACCGAGT GAGCGAATCC ATCCGCATTG GATCCAATT CTTGCCCTTG CCATCCTCT
6361 TTCTTCCCAC TTGGCGCAAC CAACACTTCC CTTGGTCTGG GTACTCGTGT TGATCTTCAC
6421 TCTCTTTTT TCTTGGCGA CCGACTTTT ATATCCGTCC TTGCTTCCCC CTGGCCGTTG
6481 TCGTTCTTC TACAACCTACC TTCCGTTCAT TATCCCTTT CTGGTTCGG TCGAGGACCC
6541 AAAAACAGAA CAATTCCGGC TCTTCCAGGT GGCTTGGGTG CGACTGTTA GCTCTTGACC
6601 ACTAGCGCT TACCTTCTCT TGATGTTTAT ATTGGATAT CATTGAACTA CTCTTCTTG
6661 AAACGGCAGA CGAACCGAAC AGTCCCTACG GTTATTAGC GATATACTGTT GTACTGATAT
6721 CCTGAGCAAG AAGAGGCAA TTATCAATT TGCACTCTCC ATCGTCGCTG CTCATCGCAG
6781 CTCCCTTGCT CGCCAATGTA TCGGCCAAC CGATTAGGAT ACCCCAACGC GATGTTCTCC
6841 GTGGTATCAA CATCACAGCA ACTTGGCGTT CGAGCACTAC CGAACCGCC CAGCGGTGGA
6901 TATGCCCTG CCGTTGTAGA CTGTCCTAAC ACCAACCGGA CGCTCCGGAA GGCGGTGGAT
6961 TTGTCGAACG AGGAGAAGAA CTGGTGTGCG ATCCGGAGGA AGAACACCAT CCAGCCCATG
7021 AGGGACCTAC TGAAGAGGGC CAACATCACT GGGTCGATT CCGAAACTTT CATGAATGAG
7081 GCCGCCAACA ACGTCTCGCA ACTGCCAAT GTGCCATTG CCATTTCTAG AGGGGGCTAT
7141 CGTCCCTCA TGAACGGCGC CGGCTTCGTT GCTGCTGCGG ATAACCGGAT TCAAAATACC
7201 ACGGGCGCAG GTGGTATTGG AGGCTTGTG CAGTCCAGCA CATATTGTA TGAAAACCA
7261 TGCCTTCTTG TGGTCTTCT TATCTCGTT TCGAGTGTCA ACTGCCAG TTGACGTTG
7321 GGCGGCTGTG GACGACCTG CTGGTGAACA TGTCTGGAC TCCATGCCCT TTTTTCTGTT
7381 CCCTAAAATC CCAAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAATTCGAG
7441 GACCGTGACT GTAAATTGCT AACGCAACTC TAGGGCCGGA CTTTCTGGTG GTGGCTGGCT
7501 TGTCGGCAGT TTGTTCTCCA ACAACTTCAG TAGCATTGAG ACCCTGCTGA GCGAGAACAA
7561 AGTCTGGGAC TTTGAGAACT CCATTTAA AGGACCAAG GAGGCTGGCC TTAGTACTGT
7621 CAACCGTATC CAGTACTGGT CCGAAGTGGC AAAGGAAGTT GCGAAGAAGA AGGATGCTGG
7681 CTTCGAGACA AGTATAACAG ACTACTGGGG CCGAGCATTG AGTTACCAAC TGATCGGAGC
7741 CGATATGGGC GGCCCGGCTT ACACCTCTC CAGCATTGCC CAGACCGACA ACTTCCAGAA
7801 GGCGAAACG CCGTCCCTA TTCTGGTAGC TGACGGCCGCG GCGCTGGAG ACACCATCAT
7861 CTCCCTCAAT GCTACCAACT ACGAGTTCAA CCCGTTCGAG ACGGGTAGCT GGACCCGAC
7921 CGTCTATGGC TTTGCGCCGA CCAAGTACCT CGGCGCCAAC TTCAGCAACG GCGTGTATCCC
7981 ATCGGGAGGC AAGTGCCTG AGGGTCTCGA CCAAGCCGGC TTGTCATGG GCACCAAGCAG
8041 CACGCTCTTC AACCAGTTCC TTTGGCCAA CATCTCCAGC TACGACGGTG TTGCCAGACG
8101 TGCTCATCGA GGCCGTGACT TCTGCTCTCA AGGAAATCGG CGCCAAGAGG ACCACGTCTC
8161 CCAAATCATC CCTAATCCGT TCCGGACTG GAACAACCGG ACCAACCCCA ACGCCGACAC
8221 GCTCGAGCTC GACCTGGTCG ACGGCGCGA AGATCTGCAG AATATCCGC TCAACCCGCT
8281 CACCAACCC GTGCGCGCCG TCGACGTCT CTTGCGCTGTC GACTCGTCCG CCGACGTGAC
8341 AAACGGCCC AATGGCACCG CCCTGCGCGC CACCTACGAG CGCACTTTCG GCTCTATTT
8401 CAACGGGACA CTCTTCCCCT CGATCCCCGA CGACTGGACG TTTATAAACC TAGGCCTCAA
8461 CAACCGCCCC TCTTCTTCG GCTGCATGT TAAGAACTTT ACCTTGAACG CCAACCAAA
8521 GGTTCCCCCCC TTAATCGTCT ATGTCCCCAA CGGCCCTAT ACACGCGCTGA GCAACGTGTC
8581 CACCTTCGAT CCGTCATACA CGATGTCTCA GCGCAACGAC ATCATGGCA ACGGATGGAA
8641 CTCAGCCACG CAGGGAAACG GCACGCTGGA TTGGAGTGG CCCACTGCG TCGCCTGCG
8701 GTTATCAGC AGGAGCTTAG ATCGGTTGGG CAGGCAGACG CCAGCCCGT GCAAGACTTG
8761 CTTGACAGG TATTGCTGGA ATGGCACAGT GAACTCCAAA GATACGGGGG TTACATGCC
8821 TGAGTTCAAG ATTGGGGATG CGCATGCCCT GGACTCGGGT GCTGTTGCTA TCGGAAAGAT
8881 GGTGAATGTC TGGTCGTCGG TTGTGGTGGG AGTTGTGGCG GCTACTTTGT TGTTGTAGGG
8941 GTAGGGGAGA CGTGATGATA TTCCAGTCTG ATGAAGTTGA GACTGGACTG GAGATCGCCA
9001 AGGATGCGGA GGGAAAGGAA TGCGTGGTGT TAATGTCATG ATGGATGAAG AGTCATGGAT
9061 CATGGAACGA CGGGCGGGGG ATATTGGATG ATGGATATAC CACACTGCAT GCATGCTCTA
9121 TTGATAGTAT GCTTGGCAT TTACGTTAA CAATCAATTG CTCCATCCTG ATGTTCTATC
9181 TTTTCGACA ATGGATTGAT ACTACTCCCTG TTGCTCGCT CTTGAGGTTG GAAGGACTTG
9241 AGGTTGGAAG GACTTGAGGT TGGTTGTTCT GAGGGAGGTT ATCGAAGTAT CATCTGTGCT

FIG. 7 continued

9301 GATGCCGATT GATAGACTGT CCTCTTCTTC GAGGCAACGA ACGGTCGGAT GAGCCTCTT
9361 AATCATGATG CTCAGTGCCA CAAAAAGGCT CCAGCACAGC TGCCCACACC TTTCTTGCCT
9421 CGCCGTTCCCT TCCTTTTCTT TTCCCCCTGT TTCCTTCTT CCTTCCATC TCATCCCGTA
9481 CCAGAGTGCC CACCGGGTAT ATATATTACC TCCTTGGCCG TTCTCCTTG ACCAATAAAAT
9541 CGCTTGGTCG AGTGGCGTAA CGGTTTACCG TCTACACTTA TCACTCAAAC CAAACCAAAC
9601 CATCGAAGAA GTGACCTATC GGTCGAGGG AACGGTGATG TTCTTACGAC CAAGTTAAC
9661 CAAAGAGCGT TCCACATCGT TGAACCGTCT CCTCCAGTTG GATCTGTTA ACTTCCGCAG
9721 CGACTGAAGA AGGTATCACT TTTTTTTGG TTCCAAAAAA AAAAAAAA ATTAC

FIG. 8

Nucleotide sequence of the *his-3 cog^E lpl* region of linkage group I in the StLawrence wild type strain of *Neurospora crassa*. This differs from that in the Lindegren strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the weak recombinator *cog^E* and also the remnant of a transposable element *Guest* within the replaceable sequence 3' of *his-3*. StLawrence strains carry *rec-2⁺* which prevents the initiation of recombination at *cog*.

1 ACCGGGAATC GTAGCGGGCG CTAAGGCCAA GCCGCGGCAC GGGTCACTGA CCCAATGCAG
61 CGCATTCCGT CAGCAACTGA AGTGGATGTA CAAGTACATA GTAGTAGATC GCAACTGGAG
121 ATCACTCGCA CCGTGCCGCA GAACAAGGGC GACGAGCCTC AGGGCAGTTT AGCCTGCCGT
181 AACAGCACAG ACCATAGCTT ATTTTCACCT GGGCGGGCGG GCGACGGCGG CACTGACATC
241 GGCAAGGCGG CATCAAGCAA CCCCTCTGTT GCTTGCCAGC TGCCGGCCAA CGTCAGCGGT
301 ACAAGGAGAA ATCTGGAAGG AAAGACTTCT GGACCCGACA GGATGGCACG CGGGAAAAGT
361 TCCCAATGCA TGAGATGAGG GGCATTTGCA TTGCCCTCCCG TCACCCAGTG CGAACCCCCAA
421 CCCCACCATA GCGTCTGTCG ATACATGGAG CGCGAAGTCG AGAAACCTGT AATTCCCTGGT
481 AACTTTCAGG TACACAGTAC GTACTGATCC TGGTATCAA CCTTGCCTGC CGAGTTTCG
541 ACGGAAAGAG GTGTGAATTG TGAAAGAGTC ATACCAAATC ACCCGATTTC CATAAAGCCC
601 GAGTCTTTTC TGTACATAAG CGACACTCGA AGCGGGCCTC ATCTTCATAG CCTGATAGCT
661 TGTAATACTC CATCCTCGTA TCTCACTGTA CCTTGAGTTC AACCCCACGT CAAACTTCAC
721 CCGACACATC GACGGATTGG GGAACAGCAC AATACCTGAA AAGCGAGAAA ACCAAACAGA
781 GGAAAACACC ATGGAGACAA CACTTCCCT CCCCTTCCTC GTCGGTGTCA GTGTTCTCC
841 CGGACTGAAT GACATCAAGG AGGGCCTCAG CGGGGAGGAA GTCTCGTGC TTGGCTGCGT
901 CTTCTTCGAG GTCAAGCCCC AGACCCTGTA GAAAATCCTG CGATTCTCA AGCGTCACAA
961 TGTCGAATTG GAGCCTACT TCGATGTAAC AGCCCTCGAG TCTATCGATG ATATTATCAC
1021 TCTCTGGAC GCCGGCGCCC GCAAGGTGTT TGTCAAGACC GAGCAGTTGG CCGACCTCTC
1081 CGCATATGGC TCCCGCGTTG CCCCCATTGT CACTGGAAGC AGCGCTGCTT TGCTTCCCTC
1141 CGCCACCGAG AGCGGCCTTT TGCTCTCCGG CCTCGATCAG ACTGCCCTCG AGGCTGCACA
1201 GTTTCTGGAG GAGGCCAGAG ACAAGAAAAT TACCCCCTTC TTCATCAAGC CCGTTCTGG
1261 GGCCGATCTC GAACAGTTCA TCCAGGTGCG CGCCAAGGCT AACGCCATCC CCATCCTGCC
1321 ATCCACTGGC TTGACAACAA AGAAGGACGA GGCCGGCAAG CTTGCCATCT CCACCATCCT
1381 CTCGAGCGTC TGGAAAGTCTG ACCGTCCCGA TGGTCTTCTC CCCACCGTTG TCGTTGATGA
1441 GCACGACACT GCTCTGGTC TGGTCTACAG CAGTGCCGAG AGTGTGAACG AGGCCCTCAG
1501 GACACAGACT GGTGTCTATC AGAGCCGGAA GCGCGGTCTC TGGTACAAGG GTGCTACTTC
1561 CGGAGACACT CAGGAGCTCG TCCGCATCTC GCTTGACTGC GATAACGATG CTCTCAAGTT
1621 TGTCGTGAAG CAGAAGGGTC GTTTCTGCCA CCTCGATCAG TCCGGCTGCT TTGGTCAGCT
1681 CAAAGGCCTT CCCAAGCTCG AGCAGACTTT GATTCGAGG AAACAGTCTG CCCCCGAGGG
1741 CTCCTACACT GCCCGTCTCT TCTCCGATGA GAAGCTAGTC CGGGCCAAGA TCATGGAGGA
1801 GGCTGAGGAG CTCTGCACCG CTCAGACCCC CCAGGAAATC GCCTTGAGG CTGCCGATCT
1861 CTTCTACTTT GCTCTTACCA GGGCGTTGC TGCCGGCGTT ACTCTGCCG ATATCGAAAG
1921 GAGCCTGAC GCCAAGAGCT GGAAGGTCAA GCGCAGGACT GGAGATGCTA AGGGTAAGTG
1981 GGCTGAGGAG GAGGGCATCA AGCCTGCGC GTCCGCTCTC GCTGCCACTT CGGCCCCCTGT
2041 CACCAAAGGAG GCCGCCAGG AGACCACCCC TGAGAAGATC ACCATGAGAC GTTTCGACGC
2101 CTCCAAGGTC TCTACCGAGG AGCTCGATGC TGCTCTCAAG CGTCTCGCG AAAAGTCGTC
2161 CGATGCCATC TACAAGATCA TTGTCCCCAT CATCGAGGAC GTCCGCAAGA ACGGCAGCAA
2221 GGCTGTTCTG TCGTACACTC ACAAGTTCGA GAAGGCTACC TCTCTTACTA GCCCCGTCTC
2281 GAAGGGCGCC TTCCCCAAGG AGCTTATGCA GCTCCCTGAG GAGACCATTG CTGCCATCGA
2341 CGTGTCTTC GAGAACATCC GCAAGTTCCA CGCCGCCAG AAGGAGGAGA AGCCCCCTCCA
2401 GGTGAGACC ATGCCCGGTG TTGTCTGCAG CGTTTCTCT CGTCCCCATCG AGGCCGTCTG
2461 CTGCTACATC CCCGGCGGTAA CCGCCGTTCT CCCCAGCACT GCCCTTATGC TGGGTGTTCC
2521 CGCCATGGTC GCCGGCTGCA ACAAGATTGT GTTGCCTCT CCTCCCCGCG CCGACGGAAC
2581 CATCACTCCC GAGATTGTCC ACGTGCTCA CAAGGTTGGG GCCGAGTCCA TCGTGTCTGC
2641 CGGCGGTGCC CAGGCCGTAG CTGCCATGGC CTACGGCACC GAGAGCATCA CCAAGGTGCA
2701 CAAGATTCTC GGCCCCGGTA ACCAGTTCGT CACTGCTGCC AAGATGTTCG TCAGCAACGA

FIG. 8 continued

2761 CACCAACGCT GCCGTTGGTA TTGACATGCC CGCTGGCCCG TCCGAGGTGC TGGTCATGCC
2821 TGACAAGGAC GCCAACCCCCG CGTTCGTTGC CTCGGATCTC CTGTCCTCAGG CTGAGCACGG
2881 CGTTGACAGT CAGGTCACTCC TGATCGCTAT TGACCTCGAC GAGGAGCATC TTCAGGCTAT
2941 TGAGGACGAG GTTCACCGTC AGGCTACGGA GCTTCCTCGC GTCCAGATTG TCCGTGGCTC
3001 CATCGCCCCAC TCGATCACCG TGCAGGTCAA GACCGTCGAG GAGGCCATGG AGCTCAGCAA
3061 CAAGTACGCT CCTGAGCACT TGATCCTCCA GATCAAGGAG GCCGAGAAGG CTGTCGATCT
3121 TGTCACTGAAC GCCGGTAGTG TCTTCATTGG CGCCTGGACT CCTGAGTCCG TTGGCGATTA
3181 CTCTGCTGGT GTTAACCACT CGCTGCGTAA GTTACATATC ATAAATAGCC CCGCTTCACA
3241 GATTCTTCTG CTAACGTCAA GACACATAGC TACCTATGGC TTTGGCAAGC AGTACTCTGG
3301 CGTCAATTTC GCCTCGTTCG TCAAGCACAT TACCAAGCTCC AACTTGACTG CCGAGGGTCT
3361 CAAAAAACGTC GGCCAGGCTG TCATGCAGTT GGCTAAGGTT GAGGAGCTCG AGGCTCACAG
3421 AAGGGCGGTC AGCATCCGTC TTGAGCACAT GAGCAAGAGC AACTAAACGG AAATTCTTT
3481 CGAAGTAGCA AAAAAAAA AAAAAAACAA GAACAAAAGG ATGTAGTGGG TTGATGTATA
3541 TCTGGGTCAT TTTGGGCACA TAGAGTAATG ATAACGAGTT TTGGACATTG TACTGTTCTG
3601 TACAGGCTGA AGATCAGTAC ATGAATCTGT TGGTAAGTGT GGAGACCCAA ACGTCCCTTG
3661 AGTTTTCTC CCTATTCCAG AGGTGCTCGT CCCTGGGTGT TTATTTCAT TATTACATCA
3721 ACCTTTTTT TTTTTTTTT TTTTCAGAT CATGGTACA TGAACGGGG AAGCACAGAC
3781 GATCGAAACG TGGATGTCAC AATGTCGCTG CAGTGATGCT GCATTGCGATG AAGCGCCCAT
3841 CTCAAATATAC TTGCACTCTT GCACGTTGCA TGTGAACCTC CCAAACAAAC GAATAAAAGA
3901 CGCGAAAAAA TGAAGATAAA AAAAAACCAT AAAAAAAATC AGAGGGAGTG TGGGAAATGG
3961 TGTCTTTAG CATTCAAGACC CCATAGCCGT GCACGCCCG GCACAGACAG GTTCATCGAT
4021 GTTGACATTG ACTGGGACAC CAGGTCTATC TATTTTATCT CCTGTCCTCT ACCATACATC
4081 GGGACATCGG ACATCTTGCT GTACCCCCCA CACCCACAAA GCCTTATAAA AGGCCACAC
4141 CCGAGGAGGT TCGGTCGGCC CCACGAACCTC TGTGCCTCCC TGCCTGTTA CAGGGACCGA
4201 ACGCTGGAGA ATCTTAATAG TTTCCTGACA TCCGGCCTAC CCGAGCAGGA AAAGGGACAG
4261 CTCATAGGGC AGGAGGGATT TGAAGATGGG AACATTTGG GTGATTGAG AGGAGGAAC
4321 AGGTACTGCA TCATGATAGT TCGGGGCAGC ATCTTGGCTG GGACATTGTT AATACCTCGA
4381 TATGATGAAG TAGGAGGGAG TTTTGCCTG TCTGCCGAA GTCCAGAGAT CTGTTTATT
4441 TTATTTTTA TGGATGTAGT GTATCAACAC CCAAGATTG GAGAATAGTA CTAGGATTG
4501 CATTTACAAG TGGAAAGTCTT GAGAATCGTT GTATATCCTT GTCTCCTCG GAATGTTAAC
4561 ATCCTACAG CGAGCGAGCG AGCGGTGGA TCGCCTGATC TGATAGGCGC AATATACGGC
4621 CGCTTCTCC GGTCTGTAG TGTAAGCTCT GTGGGCATAG TACACTAAAA AAACCTTGC
4681 ATTCATGAT CTGCCTGCTA TTCATTCCGA GCTATTTAG TGGTCACATT TCGAGGAAGA
4741 AAGAAAGCAA CTAAGATTCA CAGCCATCCA TCCATCCATA TGGAGAATA ATCCATTCCC
4801 ATGTTCCCTC CCCCCCACTA TGATGTGAC CACACGCTGC TGTCAAGATG CCAACGGTCT
4861 CAGGTACCCCT CGTCCGACTG TTTGGCATGG AGTTACATAC ACTACTAGTG TAGCCCCGGG
4921 CCAAGCTACC CCGTCAAATC TATACATATC TATAACGGGT TTCAGGGTT TCGTTCGCTG
4981 TCAATCAAGT TTGAAACATC ACTGGGCCG TTGGACGGTG TATTAGACCA TTGGCTCCCT
5041 CAGCTGTTG GCGGCTGGGC GGCTGGGTCA AACGGCAATA ACGGGACTCG AGAGGGACGA
5101 GGAGAGTCGG TTGGCTGGCT GCAATACAAG CGTCCCCACC TAACCAACGA GTCCCGTTTT
5161 CCATTGTTGTG GCCTAACCAT CATCTAGGGA TGTCAAGGGTT TGGCCGGATC AGGGTATGTT
5221 TGGTTGACTG TTGTCATGTC TGATTGGGT CATATTATGG TAGGTGTCTC GAGAACAGTA
5281 GAGTAACCGG GCCTAGCGTT TGGATGATTA CGCGAGATAT GAGTTGTGGG CCGCCATGCA
5341 GTTGCTGTG CATAAGCAGA AGTTGCTTTG GGATATATTT CTCGTCCTTC AAAGGTCACTC
5401 AGGTCTGGG ACGAACGGCA TCGCCATCCA AAGGGTTGAA CATGAGAAAC CTGAATGGCC
5461 TTTGCCTGGA AATACAAAAA GTCAAGAACAA AAATCGCTTG AGGATAGGGA CGTGGAAAGCA
5521 AGCAAATATG GTAAGAGAGG TATACATCAA CCCTGGTTCA ATTGTTAGCG TGGTTCTTCC
5581 TCCACGTCT CGTTCATGAC GGTTAACAGT ACCAGGCTAA CAATTAACACC AGGGTTGATG
5641 TGTACTGATA TGTAGGTGCT CAGCAAACCTG CCAATTCTT TGGCCCCAAG CAGCAGTTG
5701 CTGTCAGTGC TGCTCGTGTG AGCCTTGGTA GTGGAACCTA AACTGCTAAC ACAGCGCAAG
5761 TGGCGCATGTA AAGATATTGT GGGAGGATCT GTATGGATGG ATGAGATTAC TGCTTGGTGT
5821 TGGTTGCGAG GCACTGCGGC TGTTAGGCTT TGCTGTGCC CGTTCGACGA AGAAATACGC
5881 GGAACATATAA ATTGGATACC TAGACTTACT GCCTATGGGA GGTATCTACC GACGTAGCCG
5941 ACGGATTCTA GCAACATCCC GACTTTGCTT GTAGTGTACT ATGATAGCAG CACAGTGTG

FIG. 8 continued

6001 CTCCCTGTGA GAATGGGCTC TTTTTTTTTT TCCCCCTTCC CTAGGGCGTT GACTGGACTT
6061 GCTCTATTGT TCCCAAGGTA GGTGCCCGTC ATCGATTTTC CCAAGTCTCC CGCCAGATTG
6121 TCGTCATAGT GTCATGATGA CCTCGGTCGC TGCGGCTGCG TGGTTACGGG GAGCTGGGAC
6181 CGCTAGGCCT CAGTGGTTGT GCCATTCA GC GTGGGTGTGT GGAGTAGCGG TAGAGGCGCT
6241 TGGAAAGTTGT GCTAGCGGAA ACCCTGGAAAT ATCTTCTACC CTCGATTCCCT TCTCGGGCTG
6301 CCCATGTGCT GAGGTGATGC CGGGGATCTG GCGCCAATCA TCCATTGAGG TTCCCGCAGC
6361 TCCCCGGTGC CGCGCGCGGG CGCAGTTGCT CACAGGACAC ACCTAGACGC AGGGGCACAG
6421 GGGCACCGTT TGGTGTGCAA CTGGGTACCT AGCTGTAGCA AGCACTCCAC CGTCTGTGCA
6481 ATCCCCAAT CCACGGCAGG AACTTCGCAC CGCCCGCGCA CCGAGTGAGC GAATCCATCC
6541 GCATTGGATC CCAATTCTTG CCCTTGCCAT CCTTCTTCT TCCCACCTGG CGCAACCAAC
6601 ACTTCCCTTG GTCTGGGTAC TCCTGTTGAT CCTCACTCTC TTTTTTCTT GGGCGACCGA
6661 CTTTTTATAT CCGTCCTTGC TTCCCCCTGG CCGTTGTGCGT TCTTCTACA ACTACCTCC
6721 GTTCATTATC CCCTTTCTTG GTTCGGTCGA GGACCCAAAA ACAGAACAAAT TCCGGCTCTT
6781 CCAGGGGCT TGGGTGCGAC TGTTTAGCTC TTGACCACTA GCCGCTTACCC TTCTCTTGAT
6841 GTTTTATTT GGATATCATT AAACACTCT TTCTTGAAAC GGCAGACGAA CGGAACAGTT
6901 CCTACGGTAT ATTAGCGATA TACGTTGTAC TGATATTCTG AGCAAGAAGA GGCAAATTAT
6961 CAATTATGCA TCTCCCTTGC TCGCTGCTCA TCGCAGCTCC CTTGCTGCC AATGTATCGG
7021 CCGAACCCAT TAGGATACCC CAACCGCGATG TTCTCCGTGG TATCAACATC ACAGCAACTT
7081 GCCGTCGAG CACTACCGGA TTCGCCCCAGC GGTGGATATG CCCCTGCCGT TGTAGACTGT
7141 CCCAAGACCA AGCCGACGCT CGCGAAGGCC GTGGATTGT CGAACCGAGGA GAAGAACTGG
7201 TTGTCGATCC GGAGGAAGAA CACCATCCAG CCCATGAGGG ACCTCTGAA GAGGGCCAAC
7261 ATCACTGGGT TCGATTCCGA GACATTATG AATGAGGCC CCAACAAACAT CTCGCAACTG
7321 CCCAATGTG CCAATTGCCAT TTCAGGAGGC GGCTATCGT CCCTCATGAA CGCGGCCGGC
7381 TTGTTGCTG CTGCGGATAA CGAATTCAA AATACCACGG GCGCAGGTGG TATTGGAGGC
7441 TTGTTGCACT CCAGCACATA TTTGTATGTA AAGTGGTTCT TCTTATCTG TTTTCGAGTG
7501 TCAACTGCGC CAGTTCAGAG TTGGCGGCT GTGGACGACC TTGCTGGTGA ACATGTCTT
7561 GACTCCATGC CCCTCTTGC TTTCCTCAA TCAAGAAGTC GAGGACCGTG ACCGTAATC
7621 GCTAACGCAA CTCTAGGGCC GGACTTCTG GTGGTGGCTG GCTTGTGCC AGTTGTTCT
7681 CCAACAACTT CAGCAGCATT GAGACCTGC TGAGCGAGAA CAAAGTCTGG GACTTTGAGA
7741 ACTCCATCTT TAAAGGGCC C AAGGAGGCTG GCCTTAGTAC TGTCAACCGC ATTCACT
7801 GGTCCGAAGT GGCAAGGAA GTTCCAAGA AGAAGGATGC TGGCTTCGAG ACAAGTATAA
7861 CAGACTACTG GGGCCGAGCA TTGAGTTACC AACTGATCGG AGCCGATATG GGCGGCCCGG
7921 CTTACACCTT CTCCAGCATT GCCCAGACCG ACAACTTCA GAAGGCCGAA ACGCCGTTCC
7981 CTATTCTGGT AGCTGACGGC CGCGCGCTG GAGACACCAT CATCTCCCTC AATGCTACCA
8041 ACTACGAGTT CAACCCGTT GAGACGGGT A GCTGGGACCC GACCGTCTAT GGCTTGC
8101 CGACCAAGTA CCTCGGCGCC AACTTCAGCA ACGCGTGTAT CCCATCGGGA GGCAAGTGC
8161 TTGAGGGTCT CGACCAAGCC GGCTTCGTCA TGGGCACCAG CAGCACGCTC TTCAACCAGT
8221 TCCTTTGGC CAACATCTCC AGCTACGACG GTGTTGCCCG ACGTGTCTAT CGAAGCCGTG
8281 ACTTCTGTCC TCAAGGAAAT CGCGCCAAAG AGGACGACGT CTCCCAAATC ATCCCTAATC
8341 CGTTCCGTGA CTGGAACAAC CGGACCAACC CCAACGCCGA CACGCTCGAG CTCGACCTGG
8401 TCGACGGCGG CGAAGATCTG CAGAATATTC CGCTCAACCC GCTCACCCAA CCCGTGCG
8461 CCGTGGACGT CATCTCGCT GTCGACTCGT CGCCGACGT GACAAACTGG CCAATGGCA
8521 CGGCCCTGCG AGCCACCTAC GAGCGCACTT TCGGCTCTAT TTCCAACGGG ACACCTTCC
8581 CCTCGATCCC CGACGACTGG ACGTTATAA ACCTAGGCCT CAACAACCGC CCCTCTTCT
8641 TCGGCTGCGA TGTTAAGAAC TTTACCTGAA ACGCCAACCA AAAGGTTCCC CCCTTAATCG
8701 TCTATGTCCC CAACCGGCC TATAACCGCG TGAGCAACGT GTCCACCTTC GATCCGT
8761 ACACCATGTC TCAGCGCAAC GACATCATCG GCAACGGATG GAACTCAGCC ACGCAGGGAA
8821 ACGGCACGCT GGATTCGGAG TGGCCCACCTT GCGTCGCCCTG CGCGGTTATC AGCAGGAGCT
8881 TAGATCGGTT GGGCAGGCAG ACGCCAGCCG CGTGCAAGAC TTGCTTGAG AGGTATTGCT
8941 GGAATGGCAC AGTGAACCTCA AAAGATACAG GGGTTACAT GCCTGAGTTC AAGATTGCGG
9001 ATGCGCATGC CCTGGACTCG GGTGCTGTTG CTATCGAAA GATGGTGAAT GTCTGGTC
9061 CGGTTGTGGT GGGAGTTGTG GCGGCTACTT TGTGTTGTA GGGGTAGGGG AGACGTGATG
9121 ATATTCCAGT CTGATGAAGT TGAGACTGGA CTGGAGATCG CCAAGGATGC GGAGGGAAAG
9181 GAATGCGTGG TGTTAATGTC ATGATGGATG AAGGGTCATG GATCATGGAA CGACGGGGCG

FIG. 8 continued

9241 GGGATATTGG ATGATGGATA TACCACACTG CATGCATGCT CTATTGATAA TATGCTTG
9301 CATTACGTT TAACAATCAA TTGCTCCATC CTGATGTTCT ATCTTCGAC ACTGGATTGA
9361 TACTACTCCT GTTGCTTCCC TCTTGAAGTT GGAAGGACTT GAGGTTGGAA GGACTTGAGG
9421 TTGTTTGTTC TGAGGGAGGT TATCGAAGTA TCATCTGTGC TGATGCCGAT CGATAGACTG
9481 CCCTCTTCTT CGAGGCAACG AACGGTCGGA TGAGCCTCTA ATCATGATGC TCAGTGCCAC
9541 AAAAAGGCTC CAGCACAGCT GCCCACACCT TTTTGCCCTC GTCGCTCCTT CCTTTTTTC
9601 CCCCCCTTTC TTCCCTTCCA TCTCATCCCG TACCAGAGTG CCCACCGGGT ATATATATTA
9661 CCTCCTTGGC CGTTCTCCTT TGACCAATAA ATCGCTTGTT CGAGTGGCGT AACCGTTAC
9721 CGTCTACACT TATCACTCAA ACCAAACCAA ACCATCGAAG AAGTTACCTA TCGGTTCGAG
9781 GGAACGGTGA TGTTCTTACG TTCAAGTAA CCCAAAGAGC GTTCCACATC GTTGAACCGT
9841 CTCCTCCAGT TCTTGGATCT GTTTAACCTC CGCAGCGACT GAAGAAGTAA TCACTTTTT
9901 TTTTTTGTT TCCAAAAAAA AAAAAAAAAA TTAC

FIG. 9

Construction of the components of the sequence diversification cross: Parent (variant 1) and Parent (variant 2). For convenience, plasmid sequences are shown as linear. The cross hatched region in the chromosome is dispensable. Stippled sequence in the plasmid indicate the multiple cloning site for inserting foreign DNA. Crossovers in region 1 and region 2 insert the foreign sequence to be diversified into chromosome 1 of *Neurospora crassa* adjacent to the recombination hotspot *cog*. Parent (variant 2) containing a version of the foreign sequence with multiple differences from that in parent (variant 1) is similarly constructed. Parent (variant 1) and parent (variant 2) are crossed and conversion events (stippled arrow) initiated (X) at *cog*⁺ recombine the sequence differences in variant 1 and variant 2 to form new combinations. Sequences are identical except for those that distinguish variant 1 and variant 2. *rec-2* on linkage group V permits *cog*⁺ to be active. For simplicity, genes not directly related to the diversification are omitted. See text for further details.

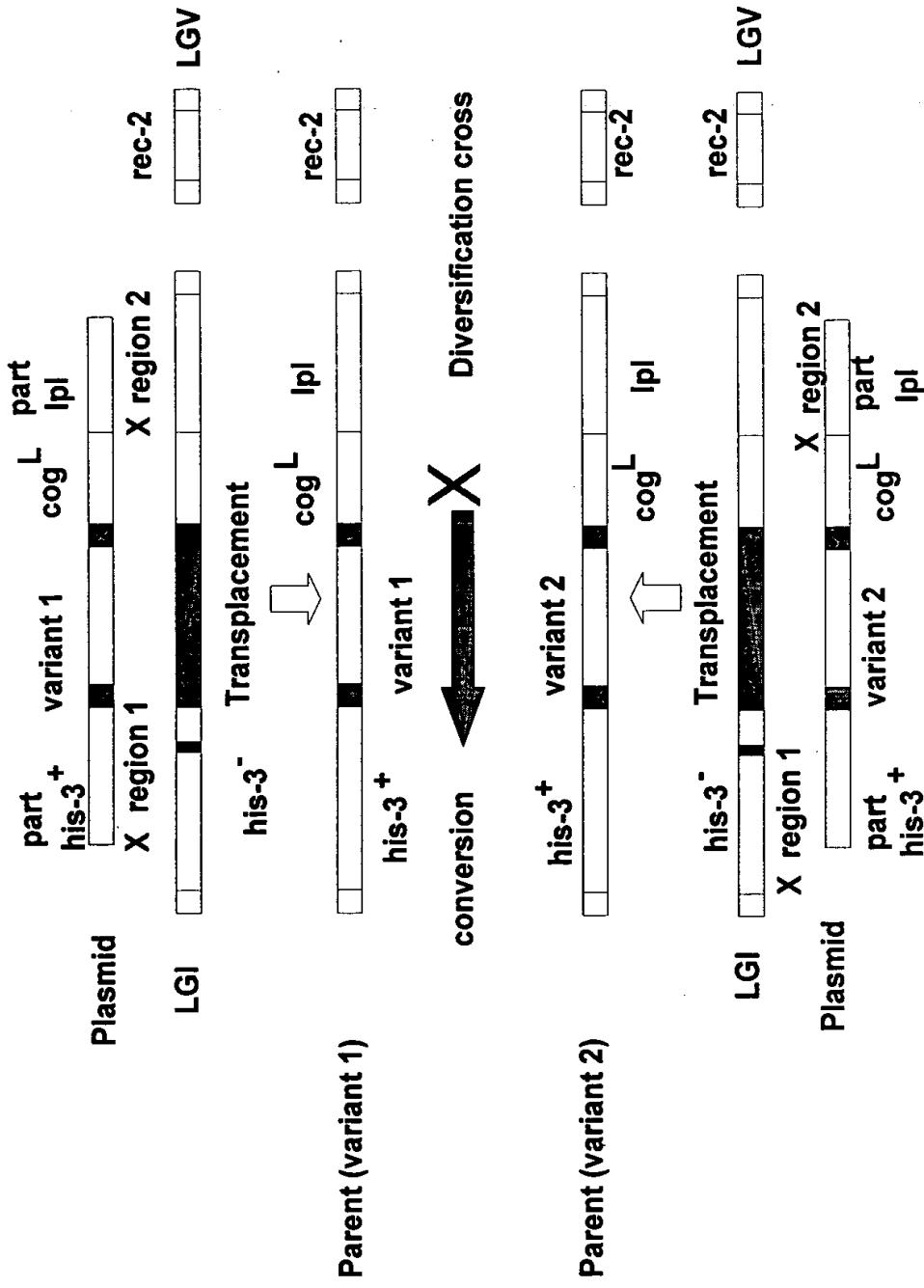


FIG. 10

Construction of parent (variant 1) and parent (variant 2) enabling selection of progeny that have experienced conversion in the foreign DNA. Complementing pairs of *his-3* alleles are used to obtain parent (variant 1) and a different pair of complementing *his-3* alleles are used to obtain parent (variant 2) as explained in the text. Parent (variant 1) and parent (variant 2) are crossed and *his⁺* recombinants are selected. These must all have experienced conversion events affecting the foreign DNA since the events begin at *cog'*. The *his-3* alleles in parent (variant 1) and parent (variant 2) are non complementing to ensure that selection yields recombinants and not aneuploid progeny having two copies of all or part of linkage group 1.

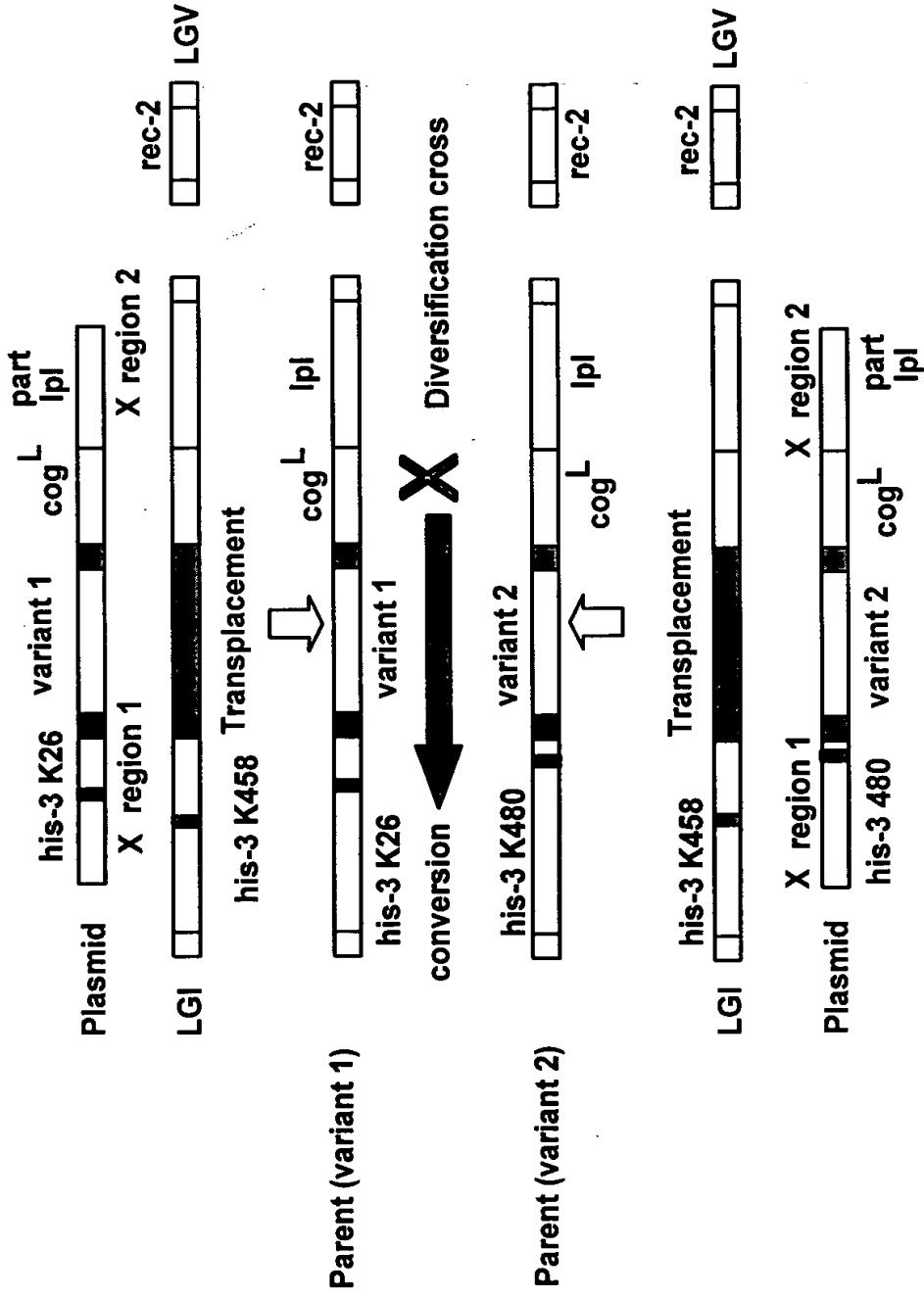


FIG. 11

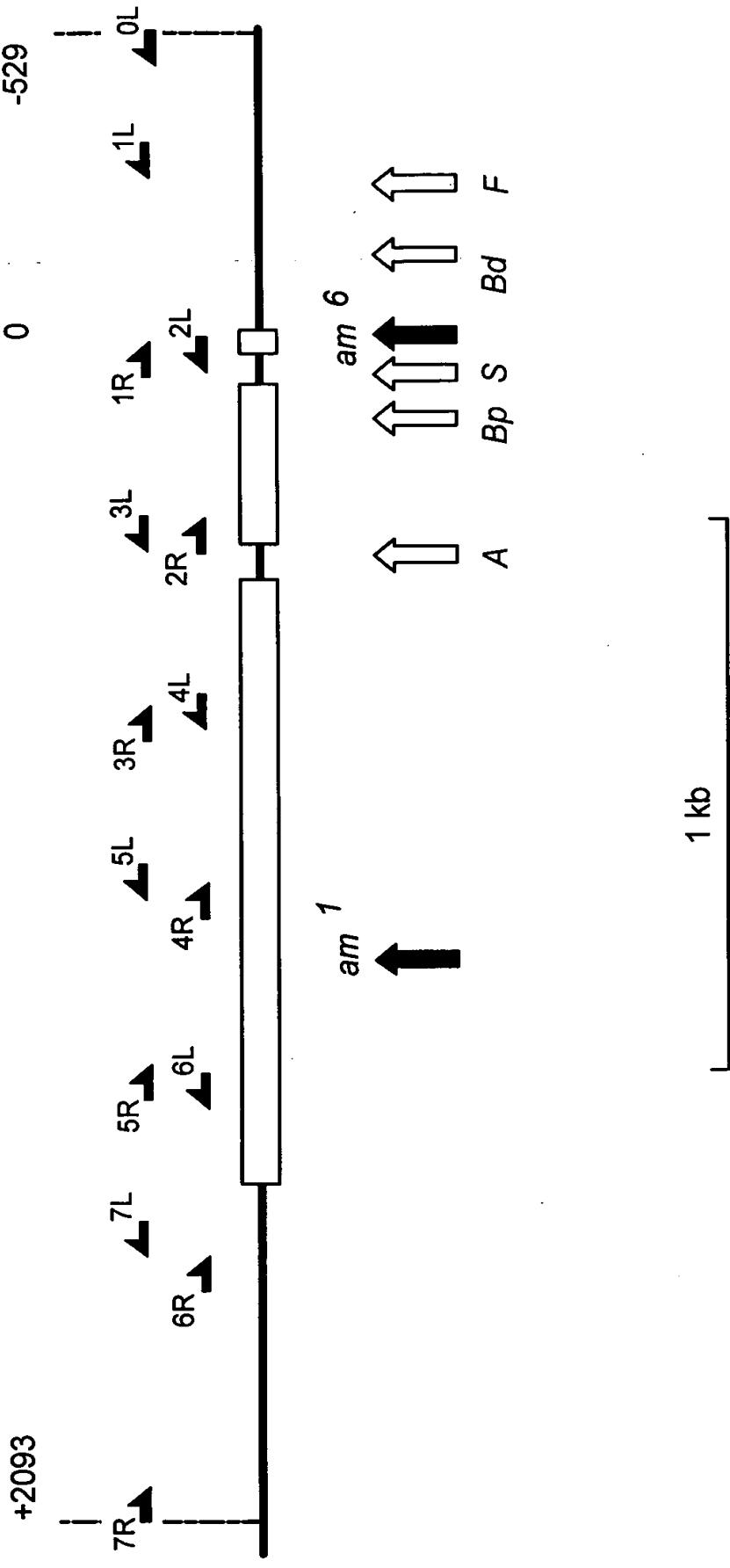


FIG. 12

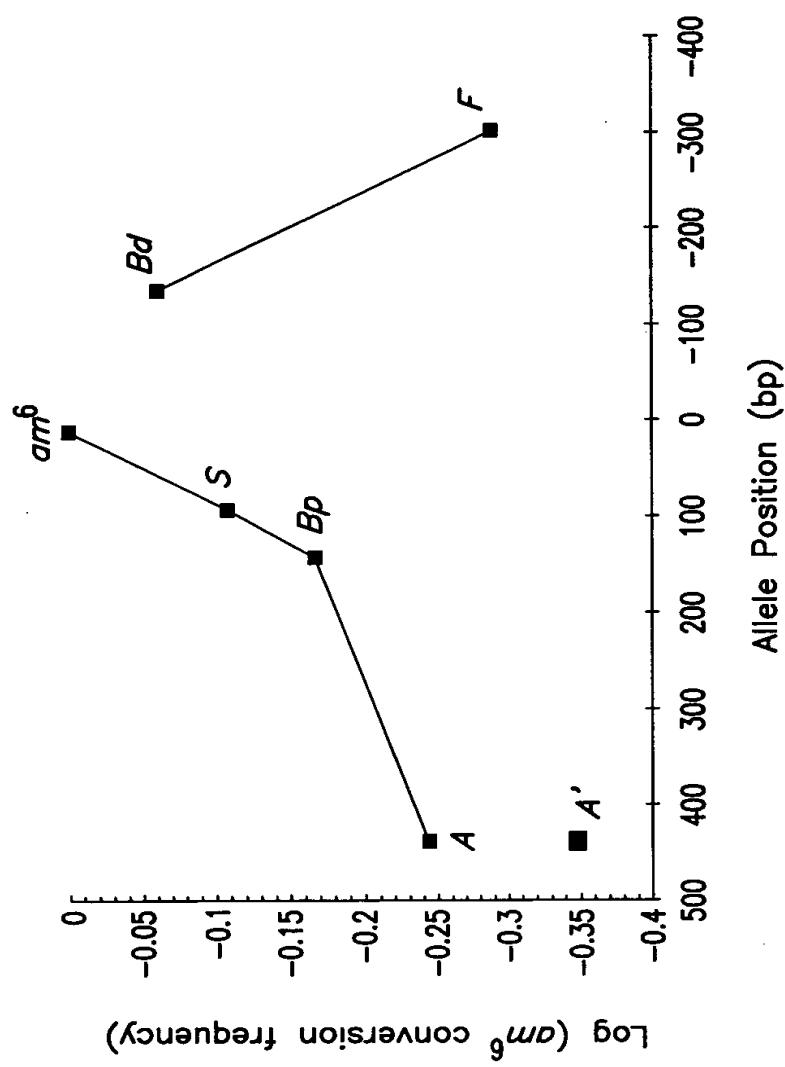


FIG. 13

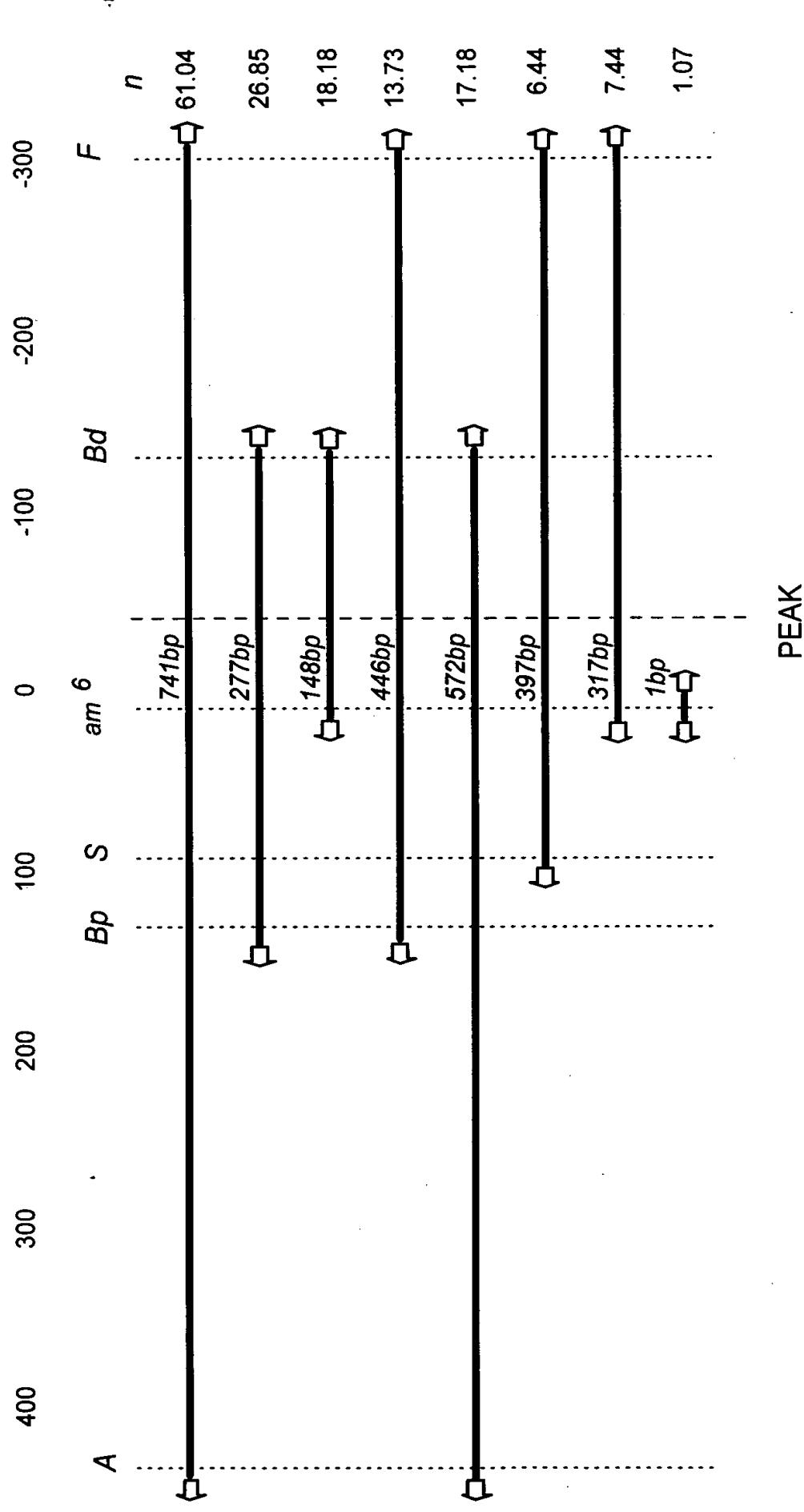


FIG. 14

